# Taking a stance: experimenting with deliberation in dialogue

**Shauna Julia Concannon**

Primary Supervisor: Professor Patrick G. T. Healey

Secondary Supervisor: Dr. Matthew Purver

Department of Electronic Engineering and Computer Science
Queen Mary University of London

Submitted in partial fulfillment of the requirements of the Degree of
*Doctor of Philosophy*

September 2017

For Christopher Bohan, patron plumber extraordinaire.

# Declaration

I, Shauna Julia Concannon, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.



Shauna Julia Concannon
September 2017

# Publications

Shauna Concannon, Pat Healey and Matthew Purver. Opening Up and Closing Down Discussion: Experimenting with Epistemic Status in Conversation. In the 39th Annual Meeting of the Cognitive Science Society (CogSci), London, July 2017.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Taking a stance: a corpus study of reported speech. In Proceedings of the 19th SemDial Workshop on the Semantics and Pragmatics of Dialogue (GoDial), Semdial Workshop, 2015a.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Shifting opinions: An experiment on agreement and disagreement in dialogue. In Proceedings of the 19th SemDial Workshop on the Semantics and Pragmatics of Dialogue (GoDial), Semdial Workshop, 2015b.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. How natural is argument in natural dialogue? In 16th Workshop on Computational Models of Natural Argument, New York, July 2016.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Opening up and closing down discussion: Experimenting with stance in conversation. In preparation.

# Acknowledgements

Firstly, I would like to thank my supervision team: Pat Healey for the many conversations, head scratching sessions and the support that he has shown throughout this process. Thank you for the endless frank, provocative and stimulating conversations, fostering renewed excitement when I had ground to a halt. I would like to thank Matthew Purver for his thoughtful advice, always delivered with clarity - from study design to detailed explanations of how to make computers do things, I've valued your contributions throughout. Graham White (together with Matt) interviewed me for the PhD programme - thank you for taking a chance on an English Literature grad and being a wonderful source of support when feeling out of place in a Computer Science Department. I would also like to thank my examiners for their thoughtful comments and incisive questions that helped me to think afresh about my work and improve it.

Countless others have supported the development of my ideas along the way. Irene, our discussions and disagreements were a source of inspiration and set in motion a train of intrigue which is still running. Quin, for being my companion in the process, proof reading, and nagging me continually to work more. And for the spare room - a good chapter. And friends who showed continuing support, accepted when I went M.I.A, and sent packages to cheer me on (Laura, Laura, Mairead, Lexie, Nikki and Vicki in particular, but all the rest of you too). And my family, who stopped asking, 'when will you get a real job', and allowed me to hermit myself away when needed and housed and fed me when that was called on too. And of course the QMUL family: Chris and Arash for showing me the ropes, teaching me how to conduct experiments, and the dark arts of the DiET; the Media & Arts Technology crew, Louis, Henrik, Dave, Nicole, Matt Jarvis, Sophie, Vincent, Louis and the rest, for all the conversations and helping me find my way, back when I was wondering what on earth a 'for loop' was.

And lastly I would like to thank Davy, who has been alongside me for much of this journey, one of the great discoveries along the way. For the ongoing conversations about research in general and particular, it's been a real pleasure.

This work was funded by EPSRC through the Media and Arts Technology Programme, an RCUK Doctoral Training Centre EP/G03723X/1.

# Abstract

How do people manage disagreements in conversation? Previous studies of dialogue have shown that the interactional consequences of disagreement are not straightforward. Although often interpreted as face-threatening when performed in an unmitigated manner, disagreement can also encourage novel contributions. This thesis explores how systematically altering the presentation of someone's stance influences the deliberative potential of a dialogue.

A corpus analysis of ordinary conversations shows that exposed disagreement occurs rarely, but that speakers can signal a potentially adversarial position in a variety of other ways. One of the most interesting among these is the way people mark their rights to speak about something. Resources such as reported speech and prefacing incongruent content with discourse markers (e.g. 'well') can be important to the management of interpersonal factors.

The idea that disagreement is problematic but also useful for deliberation is examined. Using a method that allows fine-grained manipulations of text based dialogues in real-time, agreement and disagreement fragments are inserted into a discussion dialogue. The findings show that inserting exposed disagreement violates the conventions of polite dialogue leading participants to put more effort into the production of their replies, and does not improve levels of deliberation.

This raises the question of whether manipulating apparent degrees of speaker commitment might be more important for influencing the quality of deliberation. An experiment was devised which presented oppositional content with differing degrees of 'knowingness'. The findings indicate that marking stance as knowing leads to less guarded exchanges, but does not increase deliberation. Conversely, framing statements as less knowing increases the likelihood that participants consider more alternative viewpoints, thus increasing the deliberative quality of a dialogue.

Potential applications include training guidelines for professionals developing tools to support considered debate. Implications for computational argumentation studies include the importance of interpersonal dynamics and stance construction for formulating polite arguments.

# Table of contents

# List of figures

# List of tables

# Chapter 1

# Introduction

## 1.1 Introduction

How do people manage disagreements in conversation? Disagreement is generally regarded as socially problematic and previous studies of dialogue have shown that the interactional consequences of disagreement are not straightforward. Although often interpreted as a face-threatening act when performed in an unmitigated manner, disagreement can also encourage novel contributions. Political theorist and philosophers have long acknowledged the importance of disagreement for reasoning and public opinion. Deliberative theory posits that good decisions can be reached only after multiple viewpoints have been consulted and considered (Bohman and Rehg, 1997; Dewey, 2004; Habermas, 1984). While this procedure is established and expected in formal discussion contexts, in interpersonal dialogues this process can be subject to additional social factors which can make the process more complex. In this thesis the idea that disagreement is problematic but also useful for deliberation is examined. In particular, it explores how systematically altering the presentation of someone's position on an issue affects the deliberative potential of a dialogue.

In this thesis, an empirical approach to is taken to understand the interactional effect of presenting contrasting viewpoints in dialogue. This involves interpreting how individuals form, present, adjust and respond to the expression of a speaker's position. Furthermore, dialogue is conceived as a collaborative process, in which participants actively work to establish a common ground fit for current purposes. Thus, within this conceptualisation of interaction, 'stance' is a fundamentally interactional matter, by which whatever is treated as the difference between those involved in an interaction (i.e. by the participants themselves), emerges through what is said and the display of disagreement.

In order to understand the social factors that affect this interactional phenomenon, the ways in which agreement and disagreement are presented in dialogue is explored,

with a consideration of how politeness impacts on this. In particular, it is considered whether (dis)agreements are presented more freely in situations where the risk of direct face-threat is removed, for example in reported speech, or problem solving tasks in which (dis)agreements are anticipated and perhaps welcome. Furthermore, how levels of speaker 'knowingness' affect participants' willingness to consider alternative viewpoints is also considered. Using a combination of corpus studies together with experimental investigations, this thesis explores how different methods of framing speaker position can alter the deliberative potential of a dialogue.

## 1.2 Motivation

This thesis sets out to deepen the understanding of the phenomena associated with how people present their position or viewpoint on a given topic, particularly when this is oppositional to prior contributions, together with the specific interactional consequences of such acts. The topic of disagreement is framed within a wider discourse on stance. Stance is a relatively new term in the field of linguistics and pragmatics, that is variously used in the place of prior linguistic notions such as evaluation, attitude and modality in language. However, it has been conceived and applied differently by different researchers. However, it is essentially bound to the ways in which speakers position their contribution in relation to alternatives.

A motivating factor behind this research is an interest in how individuals are led to reconsider their position on a topic, and how and when this occurs through interaction. Can the different ways of presenting a position to an interlocutor impact on the potential for deliberation within a dialogue? Disagreement is the process by which different positions on a given topic are negotiated in dialogue, however, there are a number of different ways that speakers can convey disagreement. How people present their position on a given topic, or the *epistemic stance* they take in an interaction manifests in a number of ways linguistically. Furthermore, it is essentially dialogic, i.e. relational to alternative stance positions. As such, the focus of our enquiries is not only how epistemic stance is marked in conversation, but also the articulation of agreement and disagreement, so that the trajectory of stance shifts can be tracked, and understood in terms of the interaction. In computational linguistics these challenges have been addressed to some degree in classification and detection tasks within opinion, stance and argumentation mining. More qualitative, and interactionally driven approaches have been taken by interactional linguists and conversation analysts.

## 1.3 Aim

To test these ideas we need to identify the contexts in which politeness constraints are weakened or removed. We explore how agreement and disagreement are presented in dialogue and how politeness impacts on this. In particular, we are interested in whether (dis)agreements are presented more freely in situations where the risk of direct face-threat is removed. Two examples are in reported speech, where the absence of the original addressee removes the necessity for politeness, and problem solving tasks in which (dis)agreements are anticipated and perhaps welcome. Using a combination of corpus studies together with experimental investigations, methods for detecting stance and disagreement are developed, with a view to understanding the causal effects on the deliberative quality of a dialogue.

An important corollary of this approach is it explores the consequences of disagreement, in particular, the hypothesis that it promotes shifts in *stance*. The recent literature shows an increased interest in addressing the challenge of identifying the stance a speaker holds on a given topic. There is a gap in the literature relating to the negotiation of stance in interaction and detecting when a speaker shifts their stance during the course of dialogue. While other studies have sought to extract stance and opinions from blog posts or discussion forums, the static nature of these modes of communication differ to the agenda outlined here in the following way: we are interested in how to detect stance, and in particular shifts or re-orientation of a speaker's stance, in real-time interaction. Politeness impacts on such interactions much more directly, and consequently this thesis will attempt to connect the research on stance with the literature on politeness.

### 1.3.1 Research Questions

- How do people present their position or 'point of view' on a given topic?
- Can different ways of presenting an oppositional stance affect the interactional consequences and deliberative quality of a dialogue (i.e. the number of alternatives consider before a given stance is settled upon)?
- Are certain stance constructions preferred for delivering contentious content, for example reported speech?
- How do facets of stance construction such as speaker commitment, degree of certainty and reference to an information source affect the potential for deliberation and the amount of collaboration and co-construction of stance?

## 1.4 Outline of Thesis

### 1.4.1 Chapter 2

A detailed review of the relevant literature is presented in order to situate the work conducted for this thesis. In particular, different theoretical conceptions and empirical studies on disagreement and stance in interaction are reviewed, including Appraisal Theory, Interactional Linguistics and Conversation Analysis. The conception of stance as an interactive event, as opposed to denoting of a private attitudinal state, is explored and the particular markers with which speakers can index agreement, disagreement and stance are presented.

### 1.4.2 Chapter 3

In Chapter 3 an analysis of a corpus of ordinary conversations is conducted to assess the distribution of markers indexing agreement, disagreement and 'stance' in different types of speech in the British National Corpus. Samples were drawn for comparison from reported and direct speech, as the former provides a potential context in which to investigate opinions expressed without fear of face-threat. The distributions of the markers in these different conversational contexts are explored and provide the base motivation for the experiments presented in the next two chapters.

### 1.4.3 Chapter 4

Drawing on the observations of the corpus study in Chapter 3, an experiment is presented in which the causal effects of exposed (dis)agreement is examined. Dyadic dialogues are manipulated in real-time, so that speakers' contributions appear to contain instances of exposed agreement and disagreement. As exposed disagreement rarely occurs in naturally occurring speech, this enables its effect on the dialogue and quality of the deliberation to be examined under controlled conditions.

### 1.4.4 Chapter 5

In Chapter 5 a second experiment is presented in which the apparent degree of speaker commitment to a given stance is manipulated, to present utterances as more or less 'knowing'. By investigating the impact of altering the epistemic framing of a stance as 'knowing' or 'unknowing', the interactional effect of more subtle devices for presenting oppositional content in dialogue is explored.

### 1.4.5 Chapter 6

Chapter 6 concludes the thesis and provides a discussion on the findings resulting from the work. Referring back to the relevant literature it outlines the developments made within the work and potential applications, alongside the limitations of the study and proposals for further development.

## 1.5 Associated Publications

Throughout the development of this thesis parts of the corpus analysis and experimental work have been published at international academic conferences within the fields of semantics and pragmatics and computational approaches to argumentation.

### 1.5.1 Chapter 3

The corpus study of disagreement in reported and direct speech was presented at *SemDial 2015* in Gothenburg, Sweden, in the paper Concannon et al. (2015a).

### 1.5.2 Chapter 4

The experimental work on the causal effects of exposed agreement and disagreement was also presented at *Semdial 2015* in Gothenburg, Sweden, in the paper Concannon et al. (2015b).

### 1.5.3 Chapter 5

The experimental work on the causal effects of knowing and unknowing presentations of speaker positions was submitted to *CogSci 2017* in the paper Concannon et al. (In preparation).

### 1.5.4 Additional Publications

The development of the thesis as a whole led to the publication Concannon et al. (2016), which addresses how detecting informal argument requires an understanding of interpersonal dynamics, politeness and more implicit stance markers, and the potential implications for computational approaches to argumentation in socially generated data.

## 1.6 Contributions

This thesis explores how systematically altering the presentation of someone's stance on an issue alters the deliberative potential of a dialogue. The idea that disagreement is problematic but also useful for deliberation is examined. Using a method that allows fine-grained manipulations of text based dialogues in real-time, agreement and disagreement fragments are inserted into a discussion dialogue. The causal effects of exposed disagreement has not hitherto been examined; nor has how the degree of knowingness can impact the level of constructive engagement and quality of deliberation in a dialogue.

This research sets out to extend the work on disagreement by focusing on how differing positions are negotiated in dialogue. In particular, it addresses the different ways in which oppositional positions are presented in dialogue can serve to open up or closed down a dialogue in ways that can effect the deliberative quality of a discussion. Shifts in position should signal 'constructive engagement' within a discussion, from a deliberative sense, and through this research we hope to investigate whether such an approach can help us to understand if, how and when disagreement is productive. There is currently limited research directly connecting the work from sociolinguistics on politeness and disagreement with the work on stance-taking from interactional linguistics and conversation analysis. Furthermore, there is little work evaluating the causal effect of different stance constructions through experimental methods, rather it is dominated by theoretical, corpus and fine-grained qualitative approaches.

Studies of disagreement and politeness are dominated by corpus and observational approaches. This has yielded rich insights, however it is also problematic because although it is possible to gather data on the avoidance of disagreement and the preference of politeness, it can offer little insight into what happens when disagreement does occur in dialogue. Similarly, although a number of descriptive studies based on observations of discourse marker usage exist, experimental studies in this area are still rather scarce.

Despite the limited experimental work on the interactional effects of stance marking, notable exceptions include Fox Tree (2002) which presents an experimental approach to investigate the effect of pauses and *ums* at turn exchanges, Fox Tree and Schrock (1999), which demonstrates how the discourse marker *Oh* can be used to negotiate speaker positions and Põldvere et al. (2016) which examines the effect of prosodic differences in how particular epistemic stance markers are interpreted and processed. What such an approach affords is two fold: to assess the interactional effect of certain linguistic markers in order to more fully understand their role in stance-taking and deliberative dialogue; and import the significance of interactional

dimensions of stance-taking in such a way that could be potentially utilised by computational systems. Building on such approaches, in this thesis a combination of observational and experimental methods are used to investigate the causal effects of different disagreement and stance constructions.

# Chapter 2

# Literature Review

## 2.1 Introduction

When we disagree with a conversational partner, we have a number of different ways to convey that we are taking an oppositional stance. Does how we present a stance have interactional consequences, such as opening up or closing down the dialogical landscape?

Despite being relatively familiar and intuitive concepts, establishing a clear operational understanding of what stance and disagreement are has proven difficult to pin down. The literature on disagreement and stance in dialogue incorporates qualitative and quantitative approaches and research from computational linguistics, corpus linguistics, interactional and sociolinguistics and conversation analytic traditions. The question of how stance and disagreement are enacted and the social and interpersonal consequences of such communicative acts also intersects with the concepts of politeness and 'face'.

| | |
|---|---|
| Rudy: | If it's two of you living in the house right [pause] the poll tax isn't paid by one person, it's each person is, is responsible for their own poll tax |
| Joy: | No they're not |
| Rudy: | No, if one's not working and the other one's working, the one that isn't working's supposed to get relief on it |
| Joy: | No, you're wrong, you're wrong, you're totally wrong |

Table 2.1 Exposed disagreement from the British National Corpus

| | |
|---|---|
| A: | im sure plenty of doctors think that they're on the brink |
| B: | he is very smart, so he probably isn't over exaggerating [pause] |
| B: | he might be but i think its unlikely that he is exaggerating |

Table 2.2 Subtle disagreement taken from an experiment transcript

The examples in Table 2.1 and Table 2.2 both feature disagreements, however, the way in which participants present their contributions and position them as oppositional is handled quite differently. In table 2.1 both speakers are markedly pitting themselves in opposition to one another; they are taking oppositional stances in relation to each other on the topic of council tax reduction eligibility. This is made clear through the use of turn-initial 'no' and negation, which serve to reject the prior offering, and with Joy's unequivocal assessment that Rudy is wrong, which is repeated and finally upgraded from 'you're wrong' to 'your totally wrong'. Conversely, in Table 2.2, the participants are discussing a doctor who claims to have a cure for cancer. A is refuting the validity of the doctor's claim, a judgement which B does not accept. However, rather than directly challenging or rejecting A's assessment of the situation, B offers a counter-claim, namely that doctor is smart and therefore unlikely to lie. B uses the stance marker 'I think' to frame their position and mark their assessment with evaluative language, such as 'unlikely', to convey that they are offering a considered opinion.

Individuals can also change perspective and align with their co-conversant's position. In Table 2.3, A and B are discussing who should be saved in a hypothetical scenario. Speaker A begins the extract by defending Tom, while B provides a reasoned argument for sacrificing him (that as the pilot Tom is responsible for the consequences), before re-stating their position in turn 4. A then alters their position, introducing their shift with the discourse marker 'you know what', before explicitly aligning with B's position, and upgrading from B's assertion that 'tom has to go', to 'kill tom'.

| | |
|---|---|
| A: | leave tom alone |
| A: | what |
| B: | cos hes a balloon pilot and therefore he would of known the consequences of the balloon in the first place |
| B: | no tom has to go |
| A: | you know what kill tom |
| A: | they say ignorance is bliss |

Table 2.3 Shifting positions, from an experiment transcript

The conversational excerpts in Table 2.1, Table 2.2 and table 2.3 begin to illustrate how speakers can present a viewpoint and position themselves in relation to one another, can denote disagreement in more or less obvious ways, or signal that they are altering their alignment.

'Stance' is whatever is treated as the difference between those party to an interaction, by the participants themselves, i.e. something that only emerges from the display of disagreement. It is a fundamentally interactional matter constituted by what people actually say in a particular sequential context (disagreement) and regardless of what we might say about what people 'think' or 'believe'. Thus, stance for the purposes of this thesis, is whatever is cited as being at stake in a disagreement.

To understand the ways in which differing viewpoints are negotiated in interaction and how this has been explored in the literature, this chapter will review the various theoretical characterisations of disagreement and stance from perspectives including Speech Act Theory, the Appraisal Model and more dialogical conceptions of interaction from Interactional Linguistics. The Conversation Analytic understanding of stance as an interactional achievement will then be explored, with particular reference to the Epistemic Engine outlined by Heritage (2012a). Following this, the empirical research on particular markers of disagreement and stance will be presented, ranging from the more obvious and direct markers to more subtle means of communicating disagreement in interaction.

### 2.1.1 Operationalising Disagreement and Stance

Detecting the stance a speaker holds towards a given subject can prove highly challenging, particularly in a conversational context. Research on stance detection has tended to focus on other forms of communication such as blog posts (Somasundaran and Wiebe, 2009, 2010; Walker et al., 2012a) or customer reviews, and be related to opinion mining and sentiment analysis. As Hunston (2007) observes "the phenomenon of stance represents an area of difficulty for corpus linguistics, because stance is a meaning, a type of meaning or several types of meaning, rather than a

form." While forms associated with stance have been identified, operationalising stance is not so simple a task that these forms alone can lead us to the instances in which the articulation of stance occurs. Rather, interpreting the role of stance, how, when, and in which variation it appears in dialogue, necessitates a deeper understanding of sequential aspects, context and pragmatic and semantic elements.

Disagreement poses similar challenges, in that the forms it can take are varied. While disagreement markers (like stance markers) can help us identify instances, interpreting disagreement often requires deeper understanding of the interactional context and sequential elements; inference and politeness interplay with the articulation of disagreement, adding to the difficulty of identifying and interpreting the role of disagreement in a dialogue.

There have been limited studies which directly connect research on stance in interaction with disagreement, although a number of scholars have displayed interest in both, e.g. Walker et al. (2012b,c), which deal separately with both topics from a corpus linguistics perspective. Scott (2002) investigates linguistic feature variation within disagreements using a corpus of television debate programmes. In compiling her linguistic feature set Scott (2002) incorporated research on stance (namely Biber and Finegan (1988, 1989)).

In this thesis, disagreement is conceived as interactional event in which speakers position themselves in opposition to one another, in such a way that surfaces and is addressed in the dialogue. Thus disagreement is the perceivable incongruence or misalignment in the stance positions being taken by two or more interlocutors, and it is something therefore which the speakers can identify as inconsistent between the position each respective participant takes.

'Stance' is whatever is treated as the difference between those party to an interaction, by the participants themselves, i.e. something that only emerges from the display of disagreement. It is a fundamentally interactional matter constituted by what people actually say in a particular sequential context (disagreement) and regardless of what we might say about what people 'think' or 'believe'. Thus, stance for the purposes of this thesis, is whatever is cited as being at stake in a disagreement.

This chapter begins with a contextualisation of the social consequences associated with disagreement and an overview of Politeness Theory in section 2.2.

## 2.2 Face and the social consequences of interaction

> Every person lives in a world of social encounters, involving him in face-to-face or mediated contact with other participants. In each of these contacts he tends to act out what is sometimes called a *line*-that

> is, a pattern of verbal and non-verbal acts by which he expresses his view of the situation and through this his evaluation of the participants, especially himself [...] The term *face* may be defined as the positive social value a person effectively claims for himself by the line others assume he has take during a particular contact (Goffman 1967a, p.1)

Qualitative studies show that exposed disagreement is generally avoided in conversation (Goffman, 1967a,b; Pomerantz, 1984a). One argument for the scarcity of disagreement in dialogue is anchored to the concept of politeness and the desire not be perceived as rude, nor to encroach on other's self presentation. Politeness theory has developed in response to sociological writings of Ervin Goffman on the concept of *face*. Goffman (1967a) defines face as 'the positive social value a person effectively claims for himself' through interaction and offers a model of co-operation that is enacted when an individual's face or social value is threatened during interaction. Brown and Levinson (1987), in their seminal work on politeness, explain the predisposition for the avoidance of disagreement in terms of face. Direct challenges to a speaker or disagreeing with their assertion in dialogue can constitute what is known as a *Face Threatening Act*, that is to say it can threaten the hearer's public self-image.

> When a face has been threatened, face-work must be done [...] lack of effort on the part of one person induces compensative effort from others [...] Resolution of the situation to everyone's apparent satisfaction is the first requirement [...] For example, in polite society, a handshake that perhaps should not have been extended becomes one that cannot be declined. (Goffman 1967a, pp 27-28)

This concept of *face-work* is an extremely important aspect of the rituals and behaviour that people adopt in interaction, which otherwise would be taken as indirect and arguably inefficient. Clearly, directly disagreeing with an interlocutor can pose a significant face threat. Politeness theory suggests that interlocutors would minimise disagreement to save face, employing strategic conflict avoidance techniques to mitigate the effect of any disagreement that did surface (Leech, 1980). Goffman describes a range of scenarios in which disagreements are prevented, glossed over, ignored or mitigated using avoidance tactics and corrective processes that enable a face-threat to be overcome as quickly and painlessly as possible. Thus politeness can be understood as 'strategic conflict avoidance' which 'can be measured in terms of the degree of effort put into the avoidance of a conflict situation' (Leech, 1980, p. 19). However, while Brown and Levinson continue Goffman's line of argument that

people generally cooperate in maintaining one another's face needs, at times this tendency is intentionally or unintentionally breached, result in a 'face-threating act'.

Brown and Levinson (1987) also outline three politeness super-strategies: positive politeness, negative politeness and off-record politeness, the first of which Holtgraves (1997) cites as important in disagreement. Positive politeness appeals to the positive face, and enables the disagreeing party to claim common ground and position themselves as a co-operator, thus minimising the potential threat enacted through the disagreement. Negative politeness seeks to avoid any imposition on a hearer, and maintain their freedom of action and choice. A common example is the expression of requests, which can put pressure on a hearer, and so can be formulated in such a way to minimise the negative face-threat, such as, "if you're not too busy, could you look over these papers for me?" By constructing the request in an unimposing manner, providing a possible get-out clause (I'm too busy) and through the use of the modal 'could', the speaker creates a context in which a rejection would be acceptable. Off-record politeness strategies rely entirely on inference, such as expressing it's cold, when the actual agenda is to have the window closed.

Politeness Theory supports a co-operative view of communication, in which conversation participants actively work to maintain social order by redressing and minimising any potential face-threats, an endeavour taken by all parties, through their interactions. Other perspectives on communication interpret the communicative endeavour differently. For example, Accommodation Theory posits that rather than seeking to re-align any imbalance in an interaction, interlocutors will synchronize and assume matching approaches (Giles et al., 1991). For example, if someone is agreeable their conversational partner would match them in this convivial approach, whereas if they are adopting a discursive or even combative linguistic style, then their conversational partner would be likely to adopt a similar tack and synchronicity would become more exaggerated (Giles and Smith, 1979). Accommodation Theory argues that interlocutors adopt strategies of *convergence* to integrate and identify socially with another (Giles et al., 1991); according to this theory, adopting a similar linguistic style leads to perceived communicative effectiveness (Giles and Smith, 1979) and co-operativeness (Feldman, 1968). Conversely, speech *divergence* reflects distancing from the co-conversant and can surface when confronted with perceived differences to the co-conversant. For 'constructive engagement' in discussion contexts, however, it is not really possible to converge on everything since, in the limit, this would make it impossible to actually talk about anything substantive (this will be addressed in more detail in section 2.2.2).

### 2.2.1 Minimising Disagreement

This raises the empirical question of how a 'face threat' is actually managed in conversation. Conversation Analysts have shown that when people produce assessments of situations or events, positive responses are made more quickly and clearly than negative or unaligned responses (Pomerantz, 1984a; Sacks, 1987). In natural dialogue, because of the preference to minimise disagreement and emphasise agreement, speakers often delay the delivery of *dispreferred* responses. Conversation Analysis (CA) is an approach that without introducing additional theory, looks at language used by the speakers to interpret the sequential meaning of the language. CA has shown that when people produce a response to a previous assessment, if the content is positive it is made more quickly and directly than if it is an unaligned response that might challenge the prior speaker's face. Negative or dispreferred responses are typically prefaced with a delay or an agreement token (Pomerantz, 1984a). Consequently, argumentative content can span quite a number of turns in a dialogue, and failing to consider this fully could lead to misinterpretation and false classification of stance. Disagreements can be socially problematic and so speakers often delay issuing contrasting or challenging propositions. This can be signalled through turn-initial hesitations, disfluencies and discourse markers, or by prefacing any disagreement content with an agreement.

Consider Example 2.4; in this transcription, Evaluation of an artwork, taken from (JS:I. -1) Pomerantz (1984a), participant A is inviting the others to provide their opinions on the artwork at which they are currently looking. Critical assessments are indicated in the transcript by Pomerantz with a '-' sign, while a '+' sign indicates a positive assessment. The way in which A structures their questions, 'D'yuh li:ke it?', constrains the range of appropriate responses to a polar yes/no response. D, although issuing a slight hesitation (as indicated in the transcript as 'hhh'), provides a positive appreciation in the turn directly following the initial question. Notably, this is followed by the contrastive conjunctive 'although', which initiates D's next turn, and provides some indication that they have more to add on this subject. However, it is not until some 18 turns later that D manages to contribute that they are 'not a great fan of this type of art'. In the final turn of the example D explains that that they find it reminiscent of a magazine advertisement, and state that their taste in art is more realistic. Without ever directly saying that they do not like it, it becomes clear that they don't despite having explicitly said that they do.

Subtle markers and sequential context all contribute in positioning a speaker's stance. The polar interrogative that A initially offers, leaves D with the choice of being polite, and providing the preferred response, or offering a more accurate but dispreferred response (i.e. that she doesn't like the art work), which directly

| | | |
|---|---|---|
| | A: | D'yuh li:ke it? |
| (+) | D: | .hhh Yes I do like it= (-) |
| | D: | =although I rreally::= |
| | C: | =Dju make it? |
| | A: | No We bought it, It?s a .hh a Mary Kerrida print. |
| | D: | 0:h (I k-)= |
| | A: | =Dz that make any sense to you? |
| | C: | Mn mh. I don' even know who she is. |
| | A: | She's that's, the Sister Kerrida, who, |
| | D: | Oh that's the one you to:ld me you bou:ght.= |
| | C: | Oh- |
| | A: | Ye:h |
| | D: | Ya:h. |
| | A: | Right. |
| | | (1.0) |
| | A: | It's worth something, |
| | | (1.0) |
| | A: | There's only a hundred of'm |
| | | (0.5) |
| | D: | Hmm |
| | E: | which picture is that. |
| | A: | The one that saysLife. |
| | | (1.5) |
| | A: | ( ). |
| (-) | D: | 'hhh Well I don't- I'm not a great fan of this type of a:rt. There are certain ones I see thet I like, But I like the w- + |
| | E: | =Is there ano thu way of spelling Life?. |
| (-) | D: | -more realistic-. |
| | A: | hhmh! |
| | E: | That's all I wd loo(hh)k fo(h), |
| | D: | hh! |
| (-) | D: | Yih d-know why don't got fer this type of uh: art, Becuz it- it strikes me ez being the magazine ad-verti:sement yt:pe. Which some uh-uh some a' them are really great. But tuhm I-my, taste in art is for the more uhit-t-treh- it tends tuh be realistic. |

Table 2.4 Evaluation of a new artwork from (JS:I. -1) Pomerantz (1984a)

positions her in opposition to her interlocutor. As this example highlights, offering an opinion can be significantly affected by the social factors of the interaction. If we had considered only the first two lines a different summary of the discussion would have been concluded (example 2.5):

| | |
|---|---|
| A: | D'yuh li:ke it? |
| D: | .hhh Yes I do like it= |

Table 2.5 Detail of Evaluation of a new artwork from

From example 2.5 it is possible to assume that A and D both like the painting. D's response taken in isolation could lead to erroneous analysis; if the full context of the dialogue *is* included, then a different interpretation is possible. 'Yes I do like it', is direct and seemingly unequivocal, however, when considering the full transcript and the delay preface, the dialogue reads quite differently, and the likelihood that D simply says they like it out of politeness, before providing an account for why they don't, seems much more plausible. This example highlights the ways in which individuals carefully formulate their responses to minimise potential disagreements. It also highlights the importance of paralinguistic features, such as hesitation. Before D asserts that they do like the art work they issue a breathy hesitant delay. While this may seem like noise in the data, it is actually an important indicator that D is struggling to formulate and appropriate response. Such paralinguistic content can prove vital to an accurate interpretation of the interaction.

Making and responding to assessments and assertions occurs frequently in natural dialogue. When responding to an initial assessment, an agreement may be signalled by repeating back the original assessment, but subtle details such as whether it is an exact repeat or a modified repeat can signal whether it is a strong agreement or weaker variation, modifying or downgrading the original assessment or even acting as a disagreement. Example 2.6, taken from Pomerantz (1984a), illustrates a disagreement. A pause and delay, '(hhhhh) well', is inserted, followed by a partial agreement, before the contrastive conjunctive 'but' is uttered, signalling that this is not in fact an agreement. Such mechanisms enable the speaker to take some time to formulate their disagreement, to search for a tactful way to deliver it, and to prevent the response coming across as blunt or aggressive.

Pomerantz highlights that people have a tendency to minimize disagreements. Respondents to initial assessments employ back-downs to hint at disagreement while still leaving *room to avert it*, that is, the conversant can resume with a modified assessment that may lead onto agreement. As such, there are times when honest appraisals are simply not a part of interaction: "It is not only that what would be a disagreement might not get said, but that what comes to be said may be said

A: cause those things take working at,
(2.0)
B: (hhhhh) well, they do, but
A: They aren't accidents,
B: No, they take working at, But on the other hand, some people are born with uhm (1.0) well a sense of humor, I think it's something you are born with Bea.
A: Yes. Or it's c- I have the- eh yes, I think a lotta people are, but then I think it can be developed too.

Table 2.6 Example of a disagreement from Pomerantz (1984a)

as an agreement" (Pomerantz, 1984a). In addition to hesitation, speaker B also uses the discourse marker 'well' in line 3. A turn-initial 'well' typically (but not exclusively) indicates that a disagreement is forthcoming or what follows will be in some way contrary to a prior statement (Pomerantz, 1984a). Speaker B performs an initial agreement, signalled through a turn-initial 'no' (typically regarded as a marker of disagreement) and a repeat back 'they take working at', before delivering a contrasting point of view, namely that certain traits are innate. In response, speaker A begins with a token agreement, chiming in with accord, before reverting back to their previous, contrary stance: 'I think it can be developed too'. By adding 'too' at the end of the utterance, it enables A to maintain their line of argument while conceding to the possibility that they both could be right, thus mitigating any face threat and enabling the difference of opinions to be left unresolved.

Negative or *dispreferred* responses are normally produced more slowly and are often prefaced with some form of agreement (e.g. 'Oh yes... but'); the negative assessment itself is often delayed by several turns and produced with some sort of mitigating account (Pomerantz, 1984a). Brown and Levinson (1987) also identify this preference for agreement and the occurrence of 'yes, but' constructions to mitigate the face threat associated with directly disagreeing with and interlocutor in their corpus of British data. Holtgraves (1997) confirms these observations. He coordinated discussions between pairs of participants with opposing views and analysed the resulting transcripts for positive politeness strategies outlined by Brown and Levinson (1987), as well as others that emerged through the analyses. These included: token agreement ('yes, but' constructions), hedge words, seeking common ground through the use of 'you know', developing and pursuing safe topics in which agreement could be reached, and prefacing disagreements with 'well'. This provides

some useful insights in to how disagreement is actually performed in conversation, bolstering the theoretical understandings of politeness with empirical findings.

Performing assessments and positioning one's views in relation to that of another is identified as a key component of the collaborative model of communication and central to the study of shared understanding (Goodwin, 1981; Goodwin and Goodwin, 1992). "The activity of performing an assessment is intrinsically social in that it can provide for the collaborative, but differentiated, participation of multiple actors" (Goodwin and Goodwin, 1992, p. 181). It is through assessments that participants negotiate and coordinate their perspectives with one another, providing opportunity for collaborative participation in an emerging utterance. Thus, it is through performing assessments, and agreeing, disagreeing or renegotiating the terms, that participants establish their own and others' positions and develop shared understanding. As stated by Goodwin and Goodwin (1992), "the assessment is thus not treated simply as a description, but rather as something that can be responded to, and participated in, in a special way". How participants make manifest the degree and terms of their disagreement can be conveyed through a range of subtle interactional resources. The way in which assessments are presented and the choice of response can have a significant impact on the resulting collaboration and degree of constructive engagement.

### 2.2.2 Constructive Engagement

A motivating factor behind this research is an interest in *constructive engagement* in discussion contexts. As highlighted in section 2.2, it is through the presentation of assessments that speakers negotiate their disparate and congruent perspectives on the world. Furthermore, all things being equal, there is a tendency to minimise disagreement. However, for any type of substantive discussion it is essential that differing viewpoints are consulted, and this will inevitably involve the negotiation of oppositional stances. The ways in which these procedures are managed and speaker positions presented can contribute to whether an individual may adjust the line they are taking and shift their position.

Disagreement can take a variety of forms and perform a number of different acts, which can significantly alter the illocutionary force and interactional effect of the utterance. However, in essence, disagreement is the process by which some inconsistency is treated as a difference between the parties: "Verbal disagreement is a situated activity whose function is to express an opinion (or belief) the propositional content or illocutionary force of which is – or is intended to be – partly or fully inconsistent with that of a prior (non-verbal) utterance" (Koczogh, 2013). The terminology employed to reference disagreement is wide ranging and has been

addressed in relation to argument, conflict, debate, oppositional talk and conflict talk, among others. Furthermore, aside from the appropriate terminology for discussing the presentation of oppositional stance, the way in which a turn is constructed can vary.

Recent literature on disagreement and politeness theory in Sociolinguistics and Conversation Analysis suggests that in certain contexts disagreement is appropriate (Kotthoff, 1993), can signal sociability and intimacy (Angouri and Tseliga, 2010; Schiffrin, 1984; Tannen, 1984), and rather than lead to conflict, help strengthen relationships (Georgakopoulou, 2001; Sifianou, 2012). Furthermore, in problem solving dialogues Chiu (2008) found that disagreement, when done politely, was more productive in provoking novel contributions from participants than agreement. Although face-threatening and impolite disagreement constructions are rare, in a dataset of workplace interactions deviating opinions are acceptable and unmarked disagreement is an important aspect of problem solving talk (Angouri, 2012). So, although disagreement, particularly when executed impolitely, tends to be problematic, for certain contexts, such as problem solving and discussion tasks, it can be essential in advancing the deliberative quality of a dialogue.

Thus, although there is good reason to think that disagreement ought to be socially problematic, benefits, such as encouraging novel contributions have been acknowledged (Chiu, 2008). Furthermore, Chiu (2008) also suggests that *agreement* can be potentially detrimental to a dialogue. One such context in which the surfacing of differences has been shown to be beneficial is in relation to repair, when speakers or hearers attend to some difficulty in hearing, speaking or understanding. Differences in interpretation are a basic form of disagreement and research on the phenomenon of repair shows that disruption in interaction can also be potentially beneficial to the progression of a dialogue (Colman et al., 2011; Healey, 2008), particularly if focused on the clarification of a content issue. Although instances of repair seemingly interrupt the flow of a dialogue, this attempt to address problematic talk is not necessarily negative, rather it seems to drive the conversation forward. Issuing only agreements can often lead to a lack of mutual intelligibility in fact, which is why instances of repair are so common in task-oriented dialogues (Colman et al., 2011), a context where effective co-ordination is critical to the interactional outcome. Healey (2008) demonstrates that repair processes deal directly with misalignments and have a positive effect on measures of interactional outcome. Consequently, disagreement ought to be a catalyst or precursor to a potential shift in stance, as it signals a direct challenge to a held idea, which in turn may be retained, re-negotiated or or more fundamentally re-conceived.

This, together with the findings by Chiu (2008) and Angouri (2012), suggests that disagreement can play an important role in the deliberation and problem solving

process, but *how* it is achieved can be an important factor. Whether a disagreement is interpreted as constructive engagement can depend on how it is delivered. In section 2.6.1, the specific ways that speakers convey disagreement will be addressed.

## 2.3 Forms of Disagreement

Muntigl and Turnbull (1998) analysed argument in naturally occurring dialogue between family and friends. Oppositional content was either produced quickly or in overlap, or was prefaced by agreement and pushed to the end of the turn. Examples analysed utilised negation and other negative contradiction markers or oppositional markers such as 'yes' and 'so', to accomplish the oppositional effect. Furthermore, disagreement was typically produced in four different ways in the next relevant turn after the contested stance was presented - either as irrelevancy claims, challenges, contradictions, or counter-claims Counter-claims are typically delivered with more mitigating devices, while contradiction is often achieved through negation, and is also a less direct form of disagreement. Irrelevancy claims are typically issued in a less guarded manner, in quick succession or overlap with the prior speakers contribution which is being assessed. Challenges are "often preceded by reluctance markers" and typically take interrogative question form (Muntigl and Turnbull, 1998).

68% of the disagreements analysed were issued in the next relevant turn, and were most frequently counter-claims (87 of 155; 56%), followed by contradictions (29; 18%), challenges (18; 12%), and finally irrelevancy claims (12; 8%); with contradiction and counter-claim combinations accounting for the remainder (9; 6%). As well as observing the relative frequency of the different disagreement acts, Muntigl and Turnbull (1998) observe that there are differing degrees of face-threat associated with each:

> A disagreement's degree of aggravation is likely to depend on the specific way in which an act of disagreement is done. Based on the previous research and consideration of the structural characteristics and pragmatic functions of each type of disagreement act, we propose that the ranking of disagreements from most to least face aggravating is [Irrelavancy Claim, Challenges, Contradictions, Contradictions + Counter-claims, Counter-claims].(Muntigl and Turnbull, 1998, p. 243)

Counter-claims are particularly important in substantive discussion, as it is through the presentation of new or alternative perspectives that the dialogue can progress and the deliberative quality of a dialogue develop. A number of the examples of counter-claims in Muntigl and Turnbull (1998) are framed by the propositional

marker 'I think'. It appears, that through such devices individuals introduce oppositional claims that challenge prior speakers assertions without direct negation and explicitly marking the disagreement as oppositional.

As Muntigl and Turnbull (1998) explains, a contradiction is "face-aggravating because it contains opposition markers that directly and unambiguously repudiate the other's claim. However, it is less face-aggravating than either Irrelevancy Claim or a Challenge since it does not directly attack the competency and rationality of the other speaker". However, a counter-claim "mitigates the threat to other's face by offering more information with which to negotiate the disagreement [...] It provides an alternative claim and/or reasons for why speaker disagrees, which invites negotiation of the T1 claim by opening up the topic of discussion rather than closing it down". The way in which a disagreement is constructed and the act that it performs (e.g. contraction or counter-claim, etc.) will affect the extent to which it will foster negotiation of both the self and the other's claims, and this is inextricably linked to whether it is presented in such a way that emphasises the opposition or creates space for alternative claims that are worthy of discussion.

Furthermore, disagreements are not only framed as assertions but can take the form of questions. Keisanen (2007) takes a CA approach to disagreement in conversation, and the ways in which speakers disalign themselves from prior stance positions. In particular, the process by which speakers challenge prior turns by demonstrating doubt is explored. Keisanen (2007) notes that negative yes/no interrogatives and tag questions are often used to pose challenges to the co-conversant. However, they highlight that "the negotiation of alignment does not surface in the interaction explicitly", but is available for analysis by examining the sequence organisation of relevant turns. Keisanen (2007) focuses on everyday interactions in which the challenges discussed *are* warranted "by a discrepancy in information or knowledge between the participants, and thereby relate to the social co-ordination of knowledge, rather than being presented as unanswerable or hostile assertions". These constructions, as with other interrogative forms, make a response relevant when used in conversation, i.e. they warrant some sort of response and to not respond would appear irregular. Also, because they are polar interrogatives, their grammatical form constrains the response to a yes or no response. This backs the co-conversant into a corner where they are called upon to either agree with or disagree with their partner. Heritage (2002) talks about *argumentative challenges* in news interviews and observes that negative interrogatives in turn-initial position are more assertive and less questioning than tag questions; Constructing an assessment in such a way that it invites for agreement, even though to agree would involve a concession on the part of the respondent.

To summarise, agreement and disagreement are sequentially specific forms of stance presentation, which are positioned in juxtaposition to an initial publicly stated stance. Although it is worth noting that dialogical conceptions of stance, such as Du Bois (2007); Linell and Marková (1993) suggest that stance is not only presented in reaction to what has been publicly stated in the current context, but can intersect with unspoken, imagined, or priorly issued alternate stances. Disagreement is the presentation of an oppositional stance, whereas agreement is the presentation of a stance that aligns with or supports what came before. Importantly, there are a number of different ways that oppositional stance can be presented and the interactional effect and associated face-threat vary dependent upon the way in which the opposition is constructed and presented, whether as explicit and direct, or more implicit, in the form of an unmarked counter-claim. In the next section we will address some of the different ways in which stance can be presented, which deviate from the purely exposed disagreement but can still achieve the interactional effect of signalling that the position taken is not necessarily aligned with some prior (e.g. the previous speaker's contribution).

## 2.4 Stance

The ways in which people can mark their stance is numerous; furthermore, different analytical traditions have conceived stance differently, covering a broad range of proposals. Two of the most salient aspects of stance are evaluations and assessments; the presentation of evaluations and assessments are subject to modality, which can affect the levels of certainty and speaker commitment with which they are produced. Evidentiality, providing a source for 'how I know', can also affect commitment or force of an utterance, serving to distance a speaker from a given viewpoint by presenting it as belonging to someone else, or alter the perceived certainty with which an assertion is made. For example, whether a person is detailing something they witnessed first-hand or reporting some knowledge acquired through hearsay, will result in differing levels of perceived speaker commitment.

Terms such as attitudinal stance (relating to attitudes and judgements) and alignment (the interpersonal positioning of a stance in relation to another), epistemic stance (relating to knowledge) (Glynn; Heritage, 2012a,b; Kärkkäinen, 2003; Krawczak, 2014), deonetic (relating to rights to know) affective or expressive stance (expressing feeling towards an object or event) (Lyons, 1977), writer stance (the expression of opinions in text) (Hyland, 1999), evaluative language (Hunston and Thompson, 2000) and mood and modality (Palmer, 2001), have all been used to denote related concepts.

In linguistics, the concept of stance was traditionally considered to represent subjective opinion and internalised perspectives on objects and events e.g. (Biber and Finegan, 1988, 1989; Biber et al., 1999; Conrad and Biber, 2000). Early studies in linguistics on stance related concepts such as Lyons (1977), began to create a diverse lexicon for describing the expression of stance. Palmer (2001) presents a grammatical topology of mood and modality, while Biber and Finegan (1988) address the lexical and grammatical marking of stance, with a focus on evidentiality and affect, and the role of adverbials (Biber and Finegan, 1988). Building on Halliday's conception of grammar – or 'lexicogrammar', Systemic Functional Linguists have also used the language of appraisal to explore stance, most comprehensively covered by Martin and White (2007) (c.f. Chindamo et al. (2012) for a detailed review of the different incarnations and definitions of stance in the literature).

Stance has been investigated from a formal semantics perspective in terms of 'point of view' (Mitchell, 1986); in corpus linguistics in relation to subjectivity, intersubjectivity and evaluation (Glynn; Hunston, 2007; Hunston and Thompson, 2000; Hunston et al., 2011; Krawczak, 2014); Verhagen (2005) provides an in depth study from a cognitive linguistics standpoint, concerning the intersubjectivity of language and interactional nature of stance-taking and Fitzmaurice (2004) presents a historical perspective on the development of stance markers in English. Another facet of stance expression which has been examined is the importance of certainty expression and speaker commitment or the epistemic strength of an assertion (Rubin, 2006). Non-verbal indicators of stance marking have also been explored (Prepin et al., 2013).

Langlotz and Locher (2012) explored conflictual disagreement in the comments section of MailOnline, and were interested in how people communicate *emotional stance* in online disagreements, relating their discussion to the work on impoliteness work by Culpeper (2011). They address how the public practice of dismissing a person's contribution is what can cause it to be an impolite and face threatening act. Such work highlights the politeness frameworks that surround different conversational actions, noting that those with the greatest social impact are "not such speech acts as requests and assertions, but rather challenges, and retreats, which have to do with the status of the participants, their rights and obligations, and their changing relationships in terms of social organization" (Muntigl and Turnbull, 1998, p.242).

Corpus linguists have identified linguistic features associated with epistemic stance in order to guide corpus-driven studies of subjectivity (Krawczak, 2014), or focused on evaluative language (Hunston, 2005, 2007; Hunston and Thompson, 2000). Others have focused on particular linguistic features associated with stance, such as adverbials in general (Conrad and Biber, 2000), or particular use cases such as 'surely' (Downing, 2001a). The dialogic practices of stance-taking in conversation

have been examined from an interactional linguistics perspective (Du Bois, 2007; Kärkkäinen, 2006); Field (1997) details the role of factive constructions to index epistemic stance, while Kärkkäinen (2003) analyses the function of particular epistemic stance markers, or complement taking predicates, such as 'I guess' (Karkkainen, 2007) and 'I think' (Kärkkäinen, 2003) within the frame of *epistemic scale* (Clift, 2006a; Heritage and Raymond, 2005; Kärkkäinen, 2003), for example, exploring the knowledgeable, versus ignorant, use of 'I guess' to soften claims.

A number of studies have attempted to focus on particular grammatical constructions used for stance-taking. Kärkkäinen (2003) analyses subject-verb combinations that serve as epistemic fragments to index subjectivity and stance in conversation. Scheibman (2002) studies intersubjectivity, providing an overview of frequent combinations of subjects, verb types and tenses, arguing that grammatical and lexical patterns are shaped by subjectivity and the speaker's need to personalize their contributions to discourse. Precht (2000, 2003a,b) provides a statistical analysis of stance related lexemes and grammatical constructions using a corpus of British and American English, noting cross cultural variations in the expression of stance. Myers and Lampropoulou (2012) highlight the role of impersonal pronouns in the construction of stance.

The expression of stance has also been studied across a number of mediums, for example, stance in spoken and written academic registers (Biber, 2006; Hyland, 1999, 2005), stance-taking in public discussion blogs (Myers, 2010); stance-taking in interviews (Haddington, 2004; Lampropoulou and Myers, 2012; Myers and Lampropoulou, 2012), online debate (Anand et al., 2011; Hasan and Ng, 2014; Langlotz and Locher, 2012; Somasundaran and Wiebe, 2010; Sridhar et al., 2015; Walker et al., 2012a), online diaries (Krawczak, 2014) and conversational data (Downing, 2001a; Heritage, 2008; Kärkkäinen, 2003, 2010, 2012; Kockelman, 2004; Landgrebe, 2012; Sidnell, 2012), and with a particular focus on reported speech (Clift, 2006a; Siromaa, 2012).

Sociolinguistic perspectives on stance have taken the theoretical account of stance as a dialogical act (Du Bois, 2007) and applied it to sociocultural topics such as race (Walton and Jaffe, 2011) and examined the interpersonal function of particular particles, such as 'just' (Kiesling, 2011). Work in conversation analysis takes social actions in their sequential contexts as a starting point to study how positions are adopted and negotiated in interaction. While not always labelled explicitly as relating to stance, studies focusing on agreements, disagreements and assessments (Heritage and Raymond, 2005; Pomerantz, 1984a), challenges (Keisanen, 2007), questions (Koshik, 2002), and responses (Gardner, 2001) are also relevant.

In the next section three different approaches to understanding stance will be considered in more detail, namely the Appraisal model (Martin and White, 2007), the

interactional linguistics perspective as outlined by Du Bois (2007), and the work in Conversation Analysis on stance and epistemics, as outlined by Heritage (2012a,b).

## 2.5 Different Approaches to Stance

The work on stance incorporates a number of conceptions and methodological approaches. It is possible to group the research into two broad categories: interactional or intersubjective stance and internal or subjective stance. The role of particular markers associated with stance marking in dialogue can be understood differently within these two frameworks, either as an interactionally motivated phenomenon or as the expression of personal attitude. In this thesis, the collaborative model of interaction which views interaction as co-constructed by present parties is pertained to. Consequently, in section 2.5.2 the idea of 'stance' as an interactional achievement will be considered in more detail. In the context of this thesis, it is argued that stance should only be understood in terms of the ways in which participants themselves orient to questions of similarity or difference of their positions, as manifests in the dialogue.

While the various devices and resources through which a speaker *can* take a stance have been explored, the causal effect and interactional consequences have received less attention. The specific markers associated with disagreement and marking stances as oppositional will be addressed in more detail in section 2.6, but first, the idea that stance is an interactional achievement will be considered. Accounts of modality have often interpreted hedges, downgraders, boosters and emphasisers as indicating the level of authorial confidence in "the truth of the proposition expressed" (Lyons, 1977, p. 797) (also, Palmer (2001) and Coates (1983)). Martin and White (2007) however, move the discussion of epistemic status and stance from the individualised knowledge status account, to one more explicitly attuned to interactional factors.

### 2.5.1 The Appraisal Framework

Coming from a systemic functional linguistics (SFL) perspective, Martin and White (2007) present a detailed framework for analysing the evaluative properties of a text, and the stance of a speaker or writer towards the content of an utterance and actual or potential respondents. The Appraisal framework posits that evaluative language used for the formation of a stance is the direct expression of the author's own attitude, with an additional facet oriented towards "aligning the addressee into a community of shared values and belief (Martin and White, 2007, p. 95)". There are

three components to the framework: Engagement, Attitude and Graduation. It is the first of these, Engagement, that is most central to this thesis.

Engagement involves the way in which a position is presented and the degree to which it acknowledges the availability of alternatives. A distinction is drawn between the *Monoglossic*, in which no dialogistic alternatives are acknowledged, and the *Heteroglossic* formulations that inherently acknowledge other possible alternatives. Engagement is closely related to, and often enacted through evidentiality; for example, 'John said that Jane was late', is hetroglossic, as it presents John's account of events, but permits that their could be alternatives, whereas 'Jane was late' offers a monoglossic, unequivocal account. In a sense, hetroglossic relates to a perspectival account as opposed to a factual presentation, but the degree of explicitness and the force or strength of the assertion can vary. Thus, including evidential information, for Martin and White (2007), inherently positions an assertion as one position, of multiple possible positions. However, this becomes more complicated with different evidential constructions, such as 'I know Jane was late', or 'I saw Jane arrive late', which clearly modulate the degree of alignment of the speaker with the stance presented.

Martin and White (2007), highlight that hetroglossic resources fall into two categories of intersubjective functionality: *dialogically expansive* or *dialogically contractive*. A dialogically expansive functionality better supports the deliberative endeavour, in that it fosters a communicative context in which the consideration of multiple viewpoints is supported, and in which speakers are encouraged to put forth such alternatives. They argue that writers and speakers make assessments of likelihood via modal auxiliaries, adjuncts and adverbials as well as certain mental verb/attribute projections, such as 'I think'. Conversely, 'proclaim' and 'disclaim' are listed as two resources for dialogical contraction. Disclaim includes denials, counters, which can include overt disagreement, while proclaim includes: concur, pronounce and endorse, which can encapsulate agreement as well as presenting an unequivocal factual account (e.g. the sky is blue).

The grounding of this approach in SFL means that interpersonal dimensions of an interaction are considered central to the way in which language choice and use is governed. Language in SFL is a resource that is utilised, and speakers have a choice over how to express their communicative goals, and due to the context, speaker status and environment, the probability that they will draw on particular resources is determined. Thus this is a form of macro linguistics, that attempts to inform a theory of micro linguistics, and the particular lexical, syntactical and grammatical constructions achieved through socially motivated agendas. However, analysis using the Appraisal framework, does not account for sequential aspects of interactions, nor

does it align with a collaborative model of communication, as it maintains a focus on individualised contributions, albeit with an awareness of a hearer or reader.

### 2.5.2 Interactional Stance

While Martin and White (2007) proposes an *interactionally aware* conception of stance, i.e. constructed with a reader in mind, it is in the conversation analytic work of Heritage that epistemic stance and status are positioned in terms of interactional *effect*. In essence, this means that it is through the interactive process of dialogue that stance is made manifest (Heritage, 2012b,b), a position that is similarly explored in Interactional Linguistics. Du Bois (2007) refocuses the topic of stance, on stance-taking which is a "dialogical and intersubjective activity" (Haddington, 2004), and which again permits for continual updates. As opposed to an internalised knowledge status, a purely objective knowledge state or even an interactionally aware communicative act, Heritage frames stance in terms of an epistemic engine, which is *driven* by the communicative actions and reactions of participants. For Glynn "the notions of subjectivity and epistemicity, or more specifically, how speakers (or authors) position themselves with regards to their understanding of the world and how they express this 'stance', is a central question". Glynn employs term 'epistemic stance', with 'epistemic' referring to "the individual's conception of reality" and 'stance' to "the subject's expression towards an external subject, another individual." While treatment of epistemic stance has recently come under fire as for being 'cognitivist' (Heritage), others have argued that whether or not the term is fitting, it has been used to proffer an interactionally grounded notion of stance. Consequently, a given expression in a given context will *represent* a degree of commitment to a proposition, not necessarily equate to it.

Work in the emergent field of interactional linguistics has begun to investigate stance-taking as a dialogical and fundamentally intersubjective process (Du Bois, 2007; Englebretson, 2007; Haddington, 2004). Such research posits that stance is co-constructed through interaction. This co-construction process of stance relies upon the concept of oppositions and alternatives. A stance is taken in response to an (uttered of imagined) alternative. Disagreement is one perspicuous social activity which denotes the negotiation of differing stances and a potential process by which a shift in stance can occur.

Du Bois (2007) provides a detailed theoretical approach to interactional stance, explaining that "[s]tance is a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects". Within this definition, stance is conceived as fundamentally interactional and dynamic, and con-

trary to the conception of an opinion which can be mined - which is presumed to be fixed, stance is something reflexive, that can shift throughout an interaction. Rather than separating out evaluative, affective, epistemic and other stance types, Du Bois seeks to reconcile these as part of a single unified stance act. Du Bois (2007) cites *evaluation, positioning* and *alignment* as three aspects of a unified stance act. These three aspects become important when starting to think about the linguistic features associated with stance. Evaluation refers to "the process whereby stancetaker orients to an object of stance and characterizes it as having some specific quality or value" (Du Bois, 2007, p.143), e.g. 'that's horrible'.

For Du Bois affective stance (in which speakers position themselves along an affective scale, e.g. 'I'm so glad') and epistemic stance (in which speakers present themselves as knowledgeable or ignorant, e.g. I'm not sure), both fall under what he terms *positioning*. Positioning, in line with Davies and Harré (1990), is defined as "the act of situating social actor with respect to responsibility for stance and for invoking sociocultural value"(Du Bois, 2007, p.144). Alignment, then, is the way in which the relationship between two stances is made manifest, signalled through such utterances as 'I agree', in the most explicit formulation, or a head nod, or 'yes'. Du Bois (2007) insists "a real utterance is always framed by its context of use", and that part of that context is the speaker who is responsible for it. Therefore, attempts to analyse and interpret stance acts using theoretical sentences, as has been favoured by grammarians, is limited as it cannot account for interactional motivations and social factors. As Du Bois (2007) asserts, "[s]tance is property of utterances, not of sentences, and utterances are inherently embedded in their dialogic contexts".

Linell and Marková (1993) posits that "discourse is made up of dialogical *inter-acts*, i.e. acts which are intrinsically dependent on their contextual relations, rather than of monological 'speech acts'." The essentially relational facets of dialogical interaction highlight the collaborative nature of dialogue, in which communicative acts are necessarily collective, and co-produced; dialogue is dynamic and responses are not simply responses, but also "opportunities to introduce new initiatives and new content into the interaction". Linell and Marková (1993, p. 177) continues:

> [E]ven if an utterance, such as an assertion or a request, is met with silence, that reaction, or absence of a reaction, will take on communicative significance and retroactively influence the interactional and illocutionary value of the utterance in question.

Thus, for the interactional context of substantive discussion and stance-taking, to borrow an analogy from Linell and Marková (1993) this Janus-like structure in which an utterance attends to what has come before and what may come ahead,

captures the emergent and incremental facets of stance-taking in dialogue, as well as the social sensitivity to epistemic territories.

As Linell and Marková (1993) assert: "although the discoursive acts associated with the assignments of rights and obligations are usually *not symmetrical* [...] the assignments have bilateral consequences, binding the interactants under an implicit 'contract'." A dialogic understanding of interaction is particularly pertinent for certain dialogue contexts. In relation to stance-taking, unplanned discursive dialogue is a context of heightened dialogicalism, compared to, for example a structured public debate. While public debates are prepared for, natural dialogue happens moment-by-moment; a context in which opinion formation and negotiation is manifest in the contents of the interaction. In their critique of Speech Act Theory (SAT), Linell and Marková (1993) suggest that SAT is an ill-fit for interpreting *authentic situated discourse*, due in part to its development using imagined sentences and its essentially individualistic and Cartesian conception of cognition and human agency.

### 2.5.3 Empirical evidence of stance-taking in conversation

Work in Conversation Analysis attends to this dialogical conception of interaction, but from a perspective that abstains from theoretical projections. By looking at the sequential organisation of authentic dialogue the structures and orders of social action are understood. Furthermore, it provides means for analysing such phenomena by examining the product of interaction, talk itself. So, for example, Sacks and Jefferson (1995), in discussing the significance of silences, in line with section 2.5.2, demonstrate how a silence is attributable to a speaker due to the very organisation of turn-taking and the fundamental ways in which speakers organise their contributions and select whose turn it is to speak. Stance in Conversation Analysis is viewed as an interactional achievement and something that should only be understood in the participants own terms - i.e. an oppositional stance is only recognised as such if the participants can orient to a contribution as being oppositional. The specific devices participants use and orient towards, then, are inextricably linked to the interactional consequences, and how the ensuing interaction is affected by the speakers' choices.

#### Heritage's Epistemic Engine: Knowing & unknowing

The exchange of information is clearly a key component of interaction, particularly in the domain of substantive discussion. One way that information can be exchanged is through the asking of questions and the issuing of relevant answers; however, as argued by Heritage (2012a) epistemics, and in particular epistemic imbalance between speaker is what drives contributions and sequences in a dialogue. However,

interaction is more complex than the information exchange process depicted in the Shannon and Weaver model of communication, it is not a simple transactional process, but an incremental and continually evolving play of epistemics - namely, who knows what, how they now it, their degree of certainty about and, importantly, how much of this is communicated between one conversational partner to another. Furthermore it is essentially interactive and constantly shifting as a result of joint action in this endeavour. As Paul Drew neatly summarises:

> Heritage's epistemic engine relies on participants monitoring and being continuously apprised of their epistemic status relative to their co-participants; it requires us to keep a check on what we know relative to the other, to know how we know what we know, to assess whether how we know what we know is different from how the other knows – in other words, to track quite precisely our state of knowledge with respect to some domain, relative to the other's state of knowledge. (Drew, 2012, p. 65)

Epistemic status, then, is the relative positioning in which "persons recognize one another to be more or less knowledgeable concerning some domain of knowledge as a more or less settled matter of fact". Epistemic status can be altered from moment to moment, and can be "disassembled by persons who deploy epistemic stance to appear more, or less knowledgeable than they really are" (Heritage, 2012a).

This almost performative function of epistemic stance is of particular interest. There can be social consequences of, or costs associated with seeming either knowing or unknowing (Levinson, 2012). In a general sense, Goffman (1967a) explains that "when a person volunteers a statement or message, however trivial or commonplace he commits himself and those he addresses, and in a sense places everyone present in jeopardy." More specifically, "[a] request for information positions the requester as occupying an unknowing (K-) epistemic status and the recipient in a knowing (K+) one" (Heritage, 2012a). Questions not only position the requester in an unknowing position, but bestow the obligation to respond upon the hearer (Levinson, 2012). Levinson (2012) details a question to assertion function space and observes patterns across multiple languages that suggest people prefer polar to other forms of question that require more knowledge rich responses and often disguise them as assertions. This systematic approach to questioning demonstrates an unwillingness to locate oneself in an unknowing position, nor to impose too greatly upon an interlocutor by demanding a response.

The management of rights and responsibilities is connected to participants' concerns with face (Heritage and Raymond, 2005). People can withhold questions for a range of reasons, be it fear of looking foolish or causing offence, etc., and these

'frustrated interrogative ambitions' are attributed to social or economic inhibitions according to Levinson (2012). Methods other than direct interrogation can be used to solicit an epistemic position from an interlocutor, such as hearsay: for example, compare 'are you going out tonight' and 'Tom said you'd be out tonight', one of which takes the form of a statement and the other a question, both of which create a space in the conversation for the respondent to confirm or reject the proposition that they are going out tonight. Furthermore, the use of evidentials for the epistemic framing of a statement is not only relevant when attempting to illicit information but can also be used when presenting a position. Consider the example in Table 2.7 taken from Pomerantz (1984b), in which a son tells his mother about a friend, John, who has cut his long hair into a shorter style. The son also has long hair and his mother is potentially cautious in her response, at first withholding a response and then providing as source when she does offer a stance.

| | |
|---|---|
| So: | That's *Jo*hn. He cut his *hair* by the way. |
| Mo: | Oh he did? |
| So: | Yeh |
| Mo: | Do you *like* it? |
| So: | Uh, *Yeah*, (He looks)- |
| Mo: | |
| | I heard- uh, I read two or three columns and I hear it over the TV that it's become old- it's becoming passé |

Table 2.7 Extract about Hair taken from Pomerantz (1984b)

The mother offers an assessment, a stance position that she does not claim personally - namely that long hair has fallen out of fashion. In a sense, by framing it as a report, she distances herself from the stance and assumes a less knowing position, while still contributing that position to the dialogue. Pomerantz (1984b) suggests that this strategic choice to offer a position while not openly affiliating to it may be bound to the status of the son's long hair being an 'issue', that she would not wish to directly criticize.

From this example, it is evident that making assessments, positioning a contribution as 'knowing' or 'unknowing' and providing a source or giving evidence are all key features of epistemic stance and socially complex. Furthermore, it is not inconsequential; taking an epistemic stance of 'unknowing' invites for an elaboration and projects a possible sequence expansion, whereas an interactant taking a more 'knowing' epistemic stance, creates a preference for inviting for confirmation and sequence closing; indeed 'expressions of epistemic imbalance drive sequences' (Heritage, 2012a).

### 2.5.4 Understanding stance-taking in dialogue

A key difference in the approaches outlined in this section is that Heritage, coming from the conversation analytic tradition, and Du Bois from interactional linguistics, work predominantly with spoken dialogue, whereas the Appraisal framework is developed largely in response to an imagined audience of a text: "Our framework has a prospective or anticipatory orientation in that we are concerned with the way in which the text builds for itself an audience and presents itself as engaging in various ways with this audience" (Martin and White, 2003, p. 135). Therefore, work remains to reconcile such a model when co-participants are interacting in a situated context.

Although Appraisal is concerned with the communicative effect of an utterance and challenges the monologic status of texts, due to the focus on textual resources in developing the framework, how applicable Appraisal is to dialogue is unclear. The registers used as examples in Martin and White (2007), such as newspaper articles, are written and constructed with time and thought put into augmenting the presentation of dialogic space. The full ramifications of the distinction between registers is not fully considered; the construction of 'authorial voice' in the sources which inform the framework involve premeditated construction, thus while the notion of authorial linguistic choice pertains for these contexts, to what extent this persists in real-time dialogue, in which content is produced on the fly and in response to an immediately present co-participant, is unclear.

Haddington (2004, 2005) builds upon Du Bois's theoretical framework, and integrates a conversational analytic approach that pays particular attention to the sequential context of stance in interaction. Haddington uses this integrated approach to analyse news interview data Haddington (2004, 2005). The data that Haddington is using is a rich source for exploring exposed conflicting stances and documenting the stance-taking process in dialogue. However, the deliberative aspect is somewhat missing, as news interviews are more likely to resemble polished debate format than actual discursive deliberation. The stakes in such interactions are often too high to concede a point, and lose face; consequently, the processes by which stance is renegotiated and shifts in stance occur are less likely to be exposed in this context.

In their writing on assessments, Goodwin and Goodwin (1992) stress that in conversation assessments provide for the possibility of collaborative participation in an emerging utterance. In support of this proposal they observe that feedback on assessments is often issued as the talk emerges, mid turn, offering co-participants access to influence the assessment as it is produced. Furthermore, although attending to the dialogical functions of language, the Appraisal framework doesn't address many of the specific factors associated with dialogue, such as 'face' and other aspects of a co-constructed and collaborative model of dialogue outlined in section 2.2.

Another criticism levied at the Appraisal model is its rigid conception of meaning in language as fixed, in attributing particular 'engagement' expressions as either expansive or contractive, without permitting the poly-functionality or meaning shifts (Põldvere et al., 2016). In particular Põldvere et al. (2016, p.11) demonstrates the multiple functions that complement-taking predicates can perform, with a particular focus on 'I think', by manipulating prosody and speaker status. They argue that "the discursive meanings and functions of lexical items have to be described and explained with reference to principles of meaning-making and variation" (Põldvere et al., 2016, p.11).

Consequently, the empirical research on specific markers which have been highlighted as important to how individuals present and negotiate their respective positions will be addressed in Section 2.6.

## 2.6 Linguistic Features Associated with Stance and Disagreement

In this section the specific markers relating to disagreement, stance and speaker commitment will be addressed in more detail.

### 2.6.1 Disagreement markers

The ways in which disagreement can be enacted are various. Resources such as well-prefacing (Pomerantz, 1984a), stance markers such as 'I think' (Kärkkäinen, 2003) and reported speech (Clift, 2006a; Holt and Clift, 2007) are some of the less explicit or direct ways of marking what follows as incongruous or in opposition to what went before.

Previous studies on disagreement take a distributional or corpus based approach at evidencing and analysing instances of disagreement in interaction (Abbott et al., 2011; Holtgraves, 1997; Misra and Walker, 2013; Muntigl and Turnbull, 1998; Walker et al., 2012c). These studies have provided valuable insights into the ways in which these complex social interactions are handled in different contexts, and given rise to various theories on how we process, respond to and mitigate the impact of disagreement. However, the literature also highlights that exposed disagreement rarely surfaces in naturally occurring conversation.

#### Agreement and Disagreement Cue Words

A set of more indirect indicators of agreement and disagreement are provided by cue words that are associated with agreement and disagreement but don't explicitly

formulate a turn as such. Walker et al. (2012c) analysed large datasets of forum posts to identify cue words marking features such as agreement, disagreement and sarcasm. Samples were manually annotated for levels of disagreement and agreement. In order of decreasing consensus amongst annotators the markers of disagreement were: 'really' (67% read a response beginning with this marker as prefacing a disagreement with a prior post), 'no' (66%), 'actually' (60%), 'but' (58%), 'so' (58%), and 'you mean' (57%).

These markers do not, of course, encompass all ways of doing disagreement. About 50% of respondents interpreted unmarked posts as disagreeing, highlighting the way disagreement is often enacted by more indirect means. Walker et al. (2012c) also identified markers of agreement: 'yes' (73% read a response beginning with this marker as prefacing an agreement), 'I know' (64%), 'I believe' (62%), 'I think' (61%), and 'just' (57%). Although these markers do not encompass all ways of doing disagreement. About 50% of respondents interpreted unmarked posts as disagreeing, highlighting the way disagreement is often enacted by more indirect means. Walker et al. (2012c) also identified markers of agreement: 'yes' (73% read a response beginning with this marker as prefacing an agreement), 'I know' (64%), 'I believe' (62%), 'I think' (61%), and 'just' (57%).

### 2.6.2 Explicit stance markers

There are a number of markers which explicitly index stance and communicate that what follows is intended to be interpreted as a position being taken by the speaker. The most intuitive examples would be the prefaces 'In my opinion' and 'I think'. 'I think' is the most frequently used stance marker in conversational English, despite being relatively infrequent in written data Kärkkäinen (2003). However, although it has been acknowledged as a marker which introduces personal opinions and attitudes, a number of different functions of the marker have been established. Two distinct modifying capacities have been identified, with 'I think' performing as an upgrade or downgrade modifier - a hedge or a booster. The interactional consequence of these different functions have been framed in terms of 'subjective' versus intersubjective.

Intuitively, 'I think' is the most explicit epistemic marker, in that a purely semantic interpretation would suggest that it is used to introduce an individual's thoughts on a matter, or their personal attitude towards someone or something. The interactional functionality of 'I think' has been highlighted by (Kärkkäinen, 2006), who stresses that it is not a framing device for the sharing of purely private mental states. For Martin and White (2007), the choice to include the marker (without which the ensuing content would typically still make grammatical sense) performs a dialogically expansive function. By introducing a proposition with 'I think', Martin and White

(2007) argues that it positions the utterance as one of many possible perspectives, and thus opens up the dialogic space for possible contestation. This position aligns with the interpretation of 'I think' as hedging device or downtoner.

Aijmer (2002) takes a corpus approach for the interpretation discourse markers. She acknowledges that the function of 'I think' can vary, and in particular, notes that when prefaced with 'well' it tends to serve a certain face-saving purpose, functioning as a downtoner to soften the impact of a controversial assertion. Conversely, she finds that 'now, I think' is used to introduce a subjective opinion or evaluation, and is often associated with conflict and disagreement. However, it is argued here that in both examples provided, the expression of a stance is being communicated and the face-saving efforts or explicitness of the delivery is what changes. Consequently, the interactional effect of the marker in the two examples may differ, because one is more explicitly polite, but both are delivering a potentially disagreeable stance to the hearer, with varying degrees of mitigation. Thus, while the status of 'I think' as a stance marker, or stance-taking device, is not contested, the degree of conviction with which it is issued can vary dependent on the co-location with other discourse markers or hedges, and it's sequential positioning. For example, parenthetical inclusion in a tag position is more closely associated with a hedging function (consider for example 'I think it's unfair', compared to 'it's unfair, I think'). Aijmer (1997), in her corpus, classifies all mid-turn and final position instances of 'I think', as tentative.

However, while interpretations of the function of 'I think' have tended to focus on the hedging function, Holmes (1990) contests the reductive interpretation of 'I think' as a hedging marker, stressing its dual function as both hedge and booster, dependent on syntactical positioning, prosody and contextual information relating to the participant relation, formality of the interaction and the topic under discussion. Kärkkäinen (2003) conceptualises the function of 'I think' as existing upon a scale from doubt to certainty. In light of the dual functions, Põldvere et al. (2016) challenges Martin and White (2007)'s interpretation of the engagement function of 'I think', observing that it can perform a contractive or expansive function in dialogue; such factors as speaker status, prosody and other contextual factors were found to be important in determining the perceived function (Põldvere et al., 2016).

Kärkkäinen (2003) conducted an analysis of instances of 'I think' in the Santa Barbara Corpus of spoken American English. In their data, 'I think' was found to occur in a turn-initial position in 34% of instances, turn-medial in 61%, turn-final in only 2.4% and as a separate turn in a further 2.4% of instances. By looking at the sequential positioning, Kärkkäinen (2003) found that 'I think' was often used at certain trouble spots in an interaction, specifically, when the current speaker wants "to bring in a slightly different perspective or slant to the matter expressed in the prior turn, to disagree with it, or to display uncertainty about its interactional import

or relevance" (Kärkkäinen, 2003, p. 143). Furthermore, it generally does not express a high degree of doubt or uncertainty, but falls more toward the other end of the scale.

As well as markers which directly frame an utterance as a stance, there are markers which explicitly position utterances in relation to a prior contribution. Haddington (2004) cites various linguistic resources such as intersubjective alignment markers, 'either' and 'too', that highlight how stance-taking is done in response to not only the object of conversation, but also in response to the co-conversant. Another, intuitive category is connectives, which explicitly mark relational coherence, such as 'but', 'although' or 'and'. In particular, contrastive markers, a sub-strand of discourse markers (Fraser, 1996), highlight the relational connective between speaker positions, and often signal a discordance between two oppositional stances, as demonstrated in this example, taken from Fraser (1996):

A: We can go now, children.
B: But we haven't finished our game yet.

The contrastive marker 'but' is used to introduce an oppositional stance and mark B's resistance to 'go now'. While B does not explicitly state 'I don't want to go', the utterance clearly implicates opposition, and the combination of the contrastive marker and the negation serve to convey that.

Other discourse markers have been shown to function in ways important to the presentation of speaker positions (Smith, 1998). Discourse markers are a feature of spontaneous discourse that can help speakers manage, attend to and overcome problems in speaking and understanding that can often occur when language is prepared on the fly. Although often treated as extraneous or superfluous, the addition of non-semantically critical particles, such as 'like', 'you know' and 'I mean', can have significant pragmatic effect. Fraser (1996) argues that such markers can help structure the discourse and constrain the relevance of an utterance, while Fox Tree and Schrock (1999) has demonstrated that discourse markers aid in the comprehension of dialogue. Smith (1998) highlights that discourse markers can be instrumental in how speakers negotiate their respective positions in dialogue. For example, when discussing the specific case of 'you know', Smith (1998) identifies that it is a strategic device that enables the speaker "to involve the adressee in the joint construction of a representation" and marks "statements whose implications are critical to a point being made [...] thus invit[ing] the addressee to complete the argument by drawing the appropriate inferences" (Smith, 1998, p. 196).

Experimental studies have been able to test the precise causal effects of discourse markers. For example, Fox Tree (1999) found over-hearers were able to understand and follow the content of dialogues with greater clarity than monologues (despite

thinking contrary in self-reports on performance); the significant difference between the two monologue and dialogue source content was the frequency of discourse markers and participant performance improved in correlation with the frequency of discourse markers. In trying to make sense of how discourse markers alter the interaction, Fox Tree (1999) suggests that they can contribute to a 'more friendly or creative atmosphere that fosters multiple perspectives', and that they may be more frequent in dialogues which involve the negotiation of different perspectives. For example, 'Oh' can serve to identify the interlocutors contribution as 'relevant, unanticipated and newsworthy' (Fox Tree and Schrock, 1999) – making it more polite than a mere acknowledgement, and signalling that a contribution is valued. Furthermore, it can also be used deceptively - to indicate new information has been received and a change-of-state undergone, when in fact it has not; similarly it can be withheld to signify that the information is not informative (Fox Tree and Schrock, 1999; Heritage, 1984). As Fox Tree and Schrock (1999) stress, "*Oh* does not have to indicate a genuine change of state or precede new or unanticipated information. It just has to be used to imply a change of state by the *oh* producer". Fox Tree and Schrock (1999) found that 'oh' helps listeners to integrate discourse in two ways: it primes them that what is forthcoming may involve a change of state, and it also signals that what follows should be processed separately to what came before, indicating a disjuncture between what came before and what will follow.

Detailed studies of particular markers such as 'oh' (Fox Tree and Schrock, 1999; Heritage, 1984), and 'you know' and 'I mean' (Tree and Schrock, 2002), have revealed that a more systematic effect can be determined. For example, 'you know' and 'I mean', often grouped together as similar markers that are randomly sprinkled across discourse, have been shown to have quite specific separate 'basic meanings' that relate to their interactional functions, highly relevant to negotiation dialogues. In light of a detailed review of the various use cases of the two markers, Tree and Schrock (2002, p. 744) explain:

> *You know* may be increased in dialogue because its basic meaning focusses on addressees, by inviting addressee inferences, whereas *I mean*'s basic meaning focusses on speakers, by forewarning speaker adjustments. Another way of viewing this is that *you know* encourages listeners to focus more on their own thoughts, and [...] *I mean* encourages listeners to focus more on speakers' thoughts.

This aligns with the proposed interpersonal functions of these markers, which identify their face-saving functions, either signalling shared understanding or demonstrating speaker imprecision in such a way that invites the addressee to contribute their own perspective (Holmes, 1986, 1990; Stubbe and Holmes, 1995).

Thus, a central focus of this thesis is the interactional effect of such markers, and the degree to which they foster constructive engagement and either open up or close down the dialogical landscape. For example, as described in section 2.5.3, framing an assertion as 'unknowing', invites elaboration and sequence expansion, whereas the opposite can lead to sequence closing. This process of 'dialogic expansion' and 'contraction' (Martin and White, 2003; Põldvere et al., 2016; White, 2004, 2002, 2003) can be an important factor in the deliberative quality of a dialogue. Põldvere et al. (2016) cites speaker status, prosody and co-occurrence with other stance markers as an important factor in determining the perceived function of stance markers such as 'I think', i.e. as expansive and inviting further comment, or contractive, i.e. limiting further exploration of alternate views (Põldvere et al., 2016).

In terms of interactional consequence, Kärkkäinen (2003) argues that the marker "at least weakly *projects more talk*", contrary to others who have argued for its contractive function (cf. Martin and White (2007)), in that it orients toward some problematic content or possible introduces a contrasting perspective. Furthermore, she contests the role of 'I think' as predominantly functioning as a politeness device or hedging marker, as in most use cases she examined it prefaced content issued with relatively high degrees of certainty. Rather, she explains, "*I think* generally appears in contexts that do not inherently involve a high face threat to the recipient (or speaker, for that matter)" (Kärkkäinen, 2003, p. 145). Thus, she concludes that 'I think' introduces an "explicitly personalised assessment", and is a starting point for perspective, and a boundary in the talk to orient toward some "routine interactional trouble" (Kärkkäinen, 2003, p. 146).

## 2.7 Speaker commitment: marking evidentiality and epistemicity

In this section studies which have analysed the specific mechanisms used in the presentation of a stance will be addressed. How do people *do* being knowing or unknowing? As outlined in section 2.5.3, one way in which a knowing or unknowing stance is conveyed can be through the use of interrogative format; however, a related aspect is the level of certainty a speaker demonstrates towards an assertion and this can be conveyed through epistemic markers (such as adverbials, axillary modals and hedges) and evidential markers (resources which index the source of the knowledge - such as 'I heard he moved away').

Epistemic and evidential markers are two categories of pragmatic marker that can modulate the level of perceived knowingness or commitment associated with an

assertion. Epistemic stance, can be conveyed in a number of different ways. Unlike in other languages which systematically encode epistemic stance grammatically or syntactically, in English it is marked in a variety of less formalised ways, which can be realised lexically or even prosodically (Lyons, 1977). Furthermore, many of these lexical items are pragmatic discourse markers, that invariably function in a multitude of ways. Indeed, Heritage argues that "[e]pistemic positioning is conducted through the entire resources of language and sequence organization," (Heritage, 2008, p. 309).

There are a number of linguistic devices that can index the degree of speaker commitment, or epistemicity, a person communicates about a given topic and the level of authority with which they deliver their message. For example, epistemic adverbials refer to those adverbs that address the state of the speaker or writer's knowledge and are used to express probability, possibility and certainty (Biber and Finegan, 1988). Wierzbicka (2006) identifies various types of epistemic adverbials, such as 'maybe' adverbials which perform the function of marking questionable assertions or hedging and 'surely' adverbials which express certainty.

Literature on stance and evaluative language provide a number of different linguistic features relevant to our inquiries. Willett (1988) identifies a general list of cue words that index evidentiality, while Biber and Finegan (1988), Biber et al. (1999) and Wierzbicka (2006) locate the role of particular adverbials in conveying speaker commitment. Work on hedging often incorporates some of the linguistic features outlined above, as well as additional markers, such as approximators (Prince et al., 1982; Sauerland and Stateva, 2007).

Taking a stance in response to your own prior views or those of others, is achieved through a range of linguistic features, both grammatical and lexical; often, the interpretation of these features is not monolithic but subject to the placement, prosody and other contextual factors. It has been studied from a number of different academic perspectives and the terminology employed to discuss such aspects of language are multivariate and complex. Mood, modality, evaluation, epistemic markers, evidential markers, attitudinal, propositional, positioning, footing, taking a line, stance-taking, appraisal, parentheticals, discourse particles, stance adverbials, discourse markers, modals, modifiers, boosters, hedgers, downtoners, upgraders, emphasisers are just some of the terms used to describe the range of resources through which a speaker can take a stance.[1]

[1]There is not consensus on what constitutes a stance marker and notably some terms have been used interchangeably, or even had their meaning inverted. The definitions of 'epistemic' and 'evidential' in particular have been attributed contrary definitions (e.g. Fraser (1996) uses epistemic for the source and evidential for the strength of an assertion, which is counter to most other working definitions).

In this section we will review the research which highlights the use of epistemic and evidential markers to affect the credibility or commitment to a given assertion.

### 2.7.1 Evidentiality

Expressions of evidentiality can provide reference to or encode the source of knowledge or information (e.g., the speaker/writer or someone else who may or may not be named) and the means by which the knowledge was acquired. Evidential markers, convey the source of information and reveal the basis for a proposition. In a sense they explain 'how I know'. This includes resources such as reported speech, e.g., Tracy said 'John is coming', and other attribution methods such as, 'I read that Nigel Farage is an idiot', and perception verb constructions, such as 'I saw him leave'.

Evidentiality is identified as key resource used in the presentation of stance and has been the subject of focused studies (Bergler, 2006; Biber and Finegan, 1989; Chafe, 1986; Downing, 2001b). Evidentiality also relates to the perceived epistemic strength of an assertion and has been acknowledged to contribute to the construction of authority and responsibility (Fox, 2001; Heritage and Raymond, 2005). Evidentiality is one important way that speaker commitment and accountability for an assertion can be modulated. Presenting evidence, or citing an external source for a claim can be a way to mitigate the impact of disagreement and minimizing the perceived difference between oneself and others (Pomerantz, 1984b).

Evidential markings are strategically used to accomplish social goals. For example they may function to create "distance from one's own misdeeds" (Fox, 2001). Fox (2001) suggests that evidential marking can index the social meanings of responsibility and the construction of authority, but is sensitive to context. Hunston (2007) observes that context is crucial as evaluative meaning does not occur in discrete units, but across phrases, and is cumulative, making it challenging for quantitative corpus studies.

Chafe (1986) found that although the frequencies of evidentials was not markedly different in conversation compared to academic writing, there was a difference in the types of evidentials used across context. Chafe attributes this difference to an inherent difference between speaking and writing: "A writer has much more time than a speaker to deliberate on what is being said. Speaking takes place on the fly, but a writer can mull over how best to say what is desired, and has ample time to edit what is produced [...] Speaking is an involved, social activity" Chafe (1986, p. 262). For example, belief and hearsay were more common in the conversational corpus than the academic corpus (Chafe, 1986).

Heritage and Raymond (2005) describe practices for indexing relative primacy and subordination of assessments in dialogue, asserting that when an individual

wishes to convey a lack of certainty about a claim, and thus reduce their own responsibility for accuracy of what they are saying, individuals can index this epistemic downgrade by evidential weakening. Evidential weakening is typically signally through a variety of cue words (e.g. seems, sounds). Tag questions also enable individuals to formulate their utterances as a question rather than an assertion, to defer to another speaker's epistemic rights in regards to the topic under discussion.

Biber et al. (1999), Precht (2003b) and Aikhenvald (2006) have provided categories of evidentials; however, as Chindamo et al. (2012) reflect, there is not total consensus on the name nor composition of these categories. Willet (1988) proposes a narrower sense of evidentiality, which can be picked up from the a set of cue words: hear, heard, see, saw, look, looked, sounded, sound, say, said, says, overheard, must, look/ed/s like, sound/s/ed like, seem/s/ed like, apparently, evidently, according to.

In this thesis the following categories will be adhered to: reporting verbs (e.g., said, told, reported, read); internal/inferential verbs (e.g., think, believe, feel); perceptual verbs (e.g. saw, heard, read) and relationship verbs (e.g. appears, seemed). Some markers, such as 'probably' have been classified both as hedges and epistemic adverbials in the literature. While acknowledging this duality, in this work hedges are classified as approximators and pragmatic particles which serve to make things 'more or less fuzzy' (Lakoff 1977), and separate epistemic adverbials into two categories: those which express certainty (e.g. surely, obviously) and those which express anything less than certainty, such as possibility or probability (e.g. maybe, probably). We categorise in this way, as when even expressing that something is probable, there remains room for manoeuvre, which demonstrates less authorial commitment. Furthermore, by breaking down the epistemic adverbials into those expressing certainty and those expressing uncertainty, we can learn more about the epistemic status and weakening in our corpus.

### Evidentiality: Reported Speech

Reported speech is one resource which is used for evidencing claims. Reported speech provides a novel resource for investigating the expression of stance, agreement and disagreement in dialogue. In spoken dialogue people sometimes talk about things that were said in other conversations. These instances of reported speech are typically marked by a pronoun (e.g., 'he', 'she', 'I') and an embedding verb (e.g., 'said', 'went', 'goes') followed by a rendition of the previous utterance.

> *I said, I'm not assassinating your character now but you're being very intimidating in the way that your talking to people.*[2]

[2]Theatre public meeting, September 1991, BNC-D91

*So she said, well you can't do that.* [3]

Detailed studies of the form and function of reported speech show that they are not simple verbatim reproductions of something said previously (Clark and Gerrig, 1990; Clift, 2007, 2006a; Holt, 2007, 2000). Rather, they involve the selective representation of people's own and others' conversational conduct. This allows conversational participants to use them, amongst other things, as evidence or justification for particular accounts of events, to relay complaints and disputes and to claim *epistemic priority* or privileged rights, knowledge or expertise about a topic under discussion (Clift, 2006a; Haakana, 2007; Holt, 2000; Vincent and Perrin, 1999). The non-narrative functions of reported speech have been closely associated with the expression of a *point of view*, with dialogically expansive properties (Martin and White, 2007), and argumentation, providing justification, support or authority for a particular stance (Couper-Kuhlen, 2007; Vincent and Perrin, 1999). It has been noted that direct reports are often more forthright in character, 'delivered bluntly with no mitigation' (Clift, 2007). Tannen (1984) argues that reported speech is rarely representative of what was actually said, and is rather used as a device in its own right.

Although not traditionally conceived of as evidentials, in that they do not reference an external source in support of a claim, Clift (2006a) highlights that self-reports are used for evidential effect. Furthermore, and rather than limited to story telling contexts, her analysis demonstrates that reported speech is often utilised in assessment environments.

Clift (2006a) observes the evidential capacity of reported speech to function interactionally across turns, and defines it as an *interactional evidential* compared to the stand alone evidentials. Stand alone evidentials, according to Clift (2006a) are those that only operate within the turn, such as lexical or modal means of stance marking, e.g. 'apparently', which only affect the modifies the perceived epistemic status. Interactional evidentials are positioned in response to prior assessments and "work to index the relative authority (or indeed subordination) of the speaker over a co-participant with respect to what is said" (Clift, 2006a, p. 583).

Martin and White (2007) describes reported speech as overtly dialogistic, in that it explicitly references and represents external viewpoints and utterances. Siromaa (2012) also addresses some of the ways in which direct reported speech is used in storytelling, to perform a stance-taking function; in particular a number of extracts are analysed in which reported speech is employed to convey the substance of a story which is issued in response to an initial story by their co-conversant and which, importantly, serves to reinforce the stance taken in the prior story.

[3] At home, March 1992, BNC-KCN1

### 2.7.2 Modifiers, adverbs and approximators

There are a number of linguistic devices which can make assertions more or less precise, alter the perceived speaker commitment associated with an utterance, or moderate the semantic meaning of the main clause by providing additional information which alters the degree or strength of the verb on which it operates. Emphasisers, boosters or intensifiers typically serve to upgrade the semantic meaning of an utterance, whereas downtoners or minimisers do the opposite and downgrade the strength of an assertion. Compare for example, 'I really liked it, compared to 'I quite liked it'. Pomerantz (1984a) observed that replying to an initial assessment with a partial repeat that features a downgrade, could signal disagreement. Consequently, such linguistic features can index instances where oppositional stances are presented or at odds.

The adverb 'really' in 'I really liked it', can add emphasis and boost the speaker's stance. The choice of adjective can also reveal incongruent stance positions. Consider the following example taken from Pomerantz (1984a):

A: She's a fox

B: Yeh, she's a pretty girl.

B's response uses a less emphatic adjective in response to A's fox metaphor, thus signalling a misalignment of speaker stances. As well as adjectives and adverbs such as 'quite' and 'really', *certainty adverbials*, such as 'surely', 'clearly', 'possibly' and 'probably', can also serve position speakers along a scale from certain to uncertain. Modal verbs ('can', 'could', 'may', 'might', 'shall', 'should', 'will', 'would' and 'must') and expressions affect the perceived commitment or certainty a speaker attaches to an utterance. Palmer (2001) describes epistemic modality as being "concerned with the speaker's assumptions, or assessment of possibilities, and, in most cases, it indicates the speaker's confidence or lack of confidence in the truth of the proposition expressed".

Others have advanced a more socially motivated interpretation of modals, serving to alleviate face threatening acts. Alonso-Almeida and Cruz-García (2011) highlights that modals such as 'may' and 'must' have both an epistemic and evidential function. The communicative purpose of the epistemic modal 'may' is to express probability, to avoid face threatening statements which may not be possible to justify, and can serve to protect the speaker's face and fend off external criticism (Alonso-Almeida and Cruz-García, 2011). In the context of medical texts, Vihla (1999) argues that modals and other hedges can be considered a strategy for appearing unknowing and diminish inequality between writer and reader (Vihla, 1999).

Rubin (2006) derives an empirical framework for analysing certainty about written propositions. She defines "certainty, or epistemic modality" as "a linguistic

expression of the likelihood that a particular state of affairs is, has been, or will be true". A data set of New York Times Service news reports and editorials were analysed for explicit certainty markers, factoring in the following dimensions: certainty level, perspective (the writer's point of view, a report of a direct participant, or the reported of an expert's view), focus (opinions, emotions, or judgements and facts or events), and time (past, future, etc.). She found that "central modal auxiliary verbs, gradable adjectives in their superlative degree, and adverbial intensifiers frequently express explicit certainty, while adjectival downtoners and adverbial value disjuncts rarely do so" (Rubin, 2006).

Conrad and Biber (2000) conducted a corpus study comparing the use of adverbial stance markers in conversational, academic and news registers. They employ the term 'stance' to include: epistemic stance (both in reference to the source of information and the level of commitment to a proposition), attitudinal stance (speaker's attitudes and judgements) and style stance, which relates to the manner in which information is presented (akin to discourse markers). Their analysis shows that adverbial stance markers have a variety of grammatical realizations and can occur in different clause positions, making them "among the most difficult grammatical features to analyse using computational techniques, because they are so pervasive and flexible in their distribution, and because the same forms commonly serve other grammatical functions". They observed that stance adverbials were twice as common in conversation compared to written registers. The role of stance adverbials in conversation is identified as a key resource for managing interpersonal dynamics. Conrad and Biber (2000) found that 'probably', 'actually', 'really' and 'sort of' were particularly common in conversation, and account for 70% of all epistemic markers in conversation. Through the use of such markers individuals are able to modify the content of their speech and encode important information about their commitment to a given stance and strengthen or diminish how 'knowing' they come across. This is important for face-management, and will likely have an significant impact on the appropriate responses an interlocutor can issue in response to a speaker.

### 2.7.3 Pronouns and reference devices

Pronouns and self-reference have been explored in relation to authorial responsibility in text (Hyland, 2002), medical discourse (Atkinson, 1999) and conversation (Heritage and Raymond, 2005; Lerner and Kitzinger, 2007a,b).

Hyland (2002) asserts that self-reference, typically achieved through the use of personal pronouns, is associated with commitment and knowledge claims. In addition, the first-person is effective at claiming authority when writers give information about their own stances or beliefs. This act of locating oneself can be read as a form

of self-positioning. However, it may extend beyond the self to situate anyone else who was present. Using Goffman's (1967) understanding of 'face' there is also the potential for this other-locating to jeopardize the face of others.

Oh (2007) uses a conversation analytic approach to highlight that responsibility-attribution and disagreement are often achieved through overt person reference in Korean, and Halonen (2008) demonstrates that zero-constructions in Finnish are strategically used for face-saving purposes, to distance speaker responsibility from spoken content in sensitive contexts such as group therapy. In connection to this Atkinson (1999) highlights that impersonal reportage in the passive helps in contrast to devices denoting personal agency construct "zones of responsibility" and credibility. Lerner and Kitzinger (2007a) shows that person reference relates to Goffman's notions of footing and face-work (Goffman, 1967a, 1981).

## 2.8 Language and medium

The literature drawn upon deals with the presentation of stance across a variety of spoken and written registers. It is necessary to make clear here that the concern of this thesis will be the presentation of stance in dialogue, i.e. real time exchanges that take place either face-to-face or computer mediated. The different constraints and affordances of various media affects the process of grounding (Clark and Brennan, 1991). Computer mediated real time dialogue has been acknowledged as "a hybrid language variety displaying characteristics of both oral and written language" (Ferrara et al., 1991, p.10). As the focus of this work is the linguistic expression of stance in a conversational context in which two or more co-participants are simultaneously engaged, we will draw upon both data from face-to-face interactions and co-temporal computer mediated dialogues. Consequently, the research on politeness in Computer Mediated Communication (CMC) will be briefly addressed here.

### 2.8.1 Politeness in Computer Mediated Dialogue

Early treatment of CMC suggested that the lack of social cues would negatively affect interpersonal relations (Culnan and Markus, 1987). Conversely, Walther (1992) argued that as users became more adept at processing the social information communicated via CMC over time this would become less of an issue. Walther and D'Addario (2001) cite the creative and innovative use of non-verbal cues such as emotions as evidence of emergent modes for communicating social information in CMC contexts. Park (2008) states that non-verbal communication cues were prevalent in text-based synchronous CMC channel, observing that online discourse participant expressed interpersonal and affective stance through devices such as

contraction of linguistic forms, prosodic features and typographical conventions such as the use of capital letters.

Brennan and Ohaeri (1999) conducted a study in which triads performed a memory task either face to face or via synchronous chat tool. They found that face-to-face groups hedged more than electronic groups, which they attribute to formulation costs of including hedges in typed communication. However, there was a positive correlation between word counts and hedges, suggesting that more competent typists used more hedges, confirming that the omission is related to cost of production rather than a depersonalisation affect of CMC. This, together with the equal use of questions in the two conditions, led them to conclude that a participants communicating via text still 'cared about face-management needs, and when hedging rates were low, this was because of the effort of typing.' It is also worth noting that this study was published 17 years ago, and the computer literacy rates will have improved in this time due to the widespread adoption of personal computers mobile devices and web messaging services.

So although hedging devices appear less frequently in CMC, which may give the impression of impoliteness, it is not due to a lack of regard for face management, but possibly just a by product of typing production costs. Furthermore, it is an issue which may disappear over time as individuals become more adept at typing. Alternatively, rather than just seeing a reversion to the levels of hedging witnessed in face to face communication, it may be that domain/context specific alternatives are innovated and introduced, in a similar vein to the evolution of emoticons.

Overall, the literature does not suggest that computer mediated communication instils a lack of individuation and impersonalises the identity of typists. However, the research of Park (2008) and Brennan and Ohaeri (1999), does suggest that the role of hedges in CMC can provide a useful way into understanding the dynamics of politeness in CMC.

## 2.9 Conclusion and Implications

In this chapter the devices through which speaker's present their opinions and position their contributions as oppositional were explored. From more theoretical accounts of the forms that disagreement and evaluative language can take, through to the more empirically grounded and interactional accounts, the notion that stance is an interactional matter, rather than the expression of a subjective attitudinal state was presented. In section Section 2.6 and Section 2.7 what people actually say and the particular sequential context for disagreements was addressed, ranging from subtle markers such as reported speech through to more explicit ways of stating a disagree-

ment (e.g. I disagree) or an opinion (e.g. I think X). Furthermore, the different ways of constructing disagreement and the potential interactional consequences for deliberation were considered, i.e. opening up or closing down the dialogue space.

# Chapter 3

# Taking a Stance: a Corpus Study of Reported Speech

What can reported speech tell us about the differences between how people enact disagreement and how they represent their disagreements in conversation? Which elements are preserved in the representation of (dis)agreement and which are not? To address these questions this chapter sets out to test whether there are systematic differences in the manifestation of agreement and disagreement in direct speech and reported speech in a large corpus of everyday conversations (Burnard, 1995). In particular, the distribution of markers of (dis)agreement, updates, contrast and emphasis are investigated. How speakers use these markers is compared across sub-samples of dialogue: in direct speech, in reports of their own speech and the speech of others.

## 3.1 Introduction

Qualitative studies show that people avoid exposed disagreement in conversation. This is normally attributed to politeness strategies that mitigate the face-threat involved in directly challenging or disagreeing with a conversational partner. Here we explore how agreement and disagreement are presented in reported speech, i.e. a situation where the risk of direct face-threat is removed. The difference between what is said and what is reported as said thus provides a potentially useful analytic window on the specific ways people use language to produce these different pragmatic effects. Here we focus in particular on what this contrast can tell us about the way people formulate and report on their agreements and disagreements with others.

The distribution of markers which can index agreement, disagreement, contrast and stance positioning are compared in four samples of conversational data taken from the British National Corpus (BNC). In particular, the following categories of

speech are compared: a) direct speech with reported speech b) self-reported speech ('I said') with other-reported speech ('He said', 'She said') and, in order to check effects of conversational context, c) self-reported speech with direct speech by the same speaker in their talk immediately preceding the reported speech.

This chapter proceeds by briefly setting out the Conversation Analytic (CA) research on how disagreements are typically managed in direct conversation and how reported speech can be used to present a position on a topic. Then the linguistic features that can mark disagreement, agreement, contrast and stance positioning are summarised and used to inform a quantitative analysis. This enables a comparison of the ways people both enact and report on their agreements and disagreements. Finally, the results and implications of the corpus study are presented.

### 3.1.1 Avoiding Disagreements

Making and responding to assessments and other assertions is a common feature of conversation. Conversation analysts have shown that when people produce initial assessments of situations or events, positive responses are made more quickly and clearly than negative or unaligned responses (Pomerantz, 1984a; Sacks, 1987). Negative or *dispreferred* responses are normally produced more slowly, are often prefaced with some form of agreement ('Oh yes... but') and the negative assessment itself is often delayed by several turns and produced with some sort of mitigating account (Pomerantz, 1984a).

When responding to an initial assessment, an agreement may be signalled by repeating back the original assessment, but whether this is an exact repeat or a modified repeat can signal whether it is a strong agreement or weaker variation, acting to modify or downgrade an assessment or perhaps even disagree. In the following example, taken from Pomerantz (1984a), pauses and delays, such as the '(hhhhh)', may suggest the speaker is taking some time to formulate their disagreement, or decide upon the most tactful way to deliver it:

A: cause those things take working at, (2.0)
B: (hhhhh) well, they do, but
A: They aren't accidents,
B: No, they take working at, But on the other hand, some people are born with uhm (1.0)
B: well a sense of humor, I think it's something you are born with Bea.
A: Yes. Or it's c- I have the- eh yes, I think a lotta people are, but then I think it can be developed too.

Example 1

In addition to the hesitation, speaker B also uses the discourse marker *well*, often used to highlight that a disagreement is forthcoming. Furthermore, speaker A performs an initial agreement by repeating back *they take working at*, before delivering a contrasting point of view, namely that certain traits are innate. In response speaker A also offers an appeasing agreement, before reverting back to their previous, contrary stance, *I think it can be developed too*. This small extract highlights many of the devices, such as hesitation, negation, and discourse markers, that are employed when managing disagreement in dialogue.

The CA observations highlight the ways that people normally avoid exposing disagreements directly (unless of course they intend to be abrupt or confrontational). Consequently explicit markers of disagreement should tend to be rare in conversation and much less common than explicit markers of agreement. How would we expect these phenomena to play out in reported speech?

### 3.1.2 Reported speech

Direct challenges and disagreement in conversation are socially problematic. As highlighted in the Chapter 2, exposed disagreement is generally avoided (Pomerantz, 1984a) because it is potentially face threatening (Brown and Levinson, 1987). If people are reluctant to expose disagreements directly then reported speech provides a potentially useful context in which prior disagreements could be presented more explicitly; the original addressee is absent which reduces concerns about politeness and the likelihood of a challenge to the speaker's version of events.

As discussed in section 2.7.1, reported speech involves the selective representation of people's own and others' conversational conduct and their position on a given topic.

| | |
|---|---|
| 2a) | So *she said* well you can't do that.<br>I said I bloody can do that.<br>That is my caravan, not yours.<br><pause> I said I can do it.<br>And I will do it if that man's not gonna see me.[a] |
| 2b) | I actually had phrases like bra burning thrown at me and I said that feminism to me is about having the same opportunities as men [...] everybody round the table said well yes I believe in that too<br>*I said* well then you're feminists[b] |

[a] At home, March 1992, BNC-KCN1
[b] Adapted from Informal meeting, BNC-HYY2

Table 3.1 Example 2

Reported speech can be used to express a point of view (that of another person's as in 2a) or your own as in 2b), to evidence claims or justify particular accounts of events, to relay complaints and to claim *epistemic priority* or privileged rights, knowledge or expertise about a topic under discussion (Clift, 2006a; Haakana, 2007; Holt, 2000; Vincent and Perrin, 1999). For example, in the extract from the BNC Dave produces an assessment and then quotes someone who echoes his opinion:

| | |
|---|---|
| Dave: | Yeah, but it's got worse and worse |
| Keith: | Oh right, yeah |
| Dave: | and he said it's got worse since he got the new switchboard operator, which makes you think [1] |

Table 3.2 Example 3

Direct reports are often more forthright and feature less mitigation (Clift, 2007). This unique context in which the usual constraints of politeness, i.e. the formulation of utterances is taken with particular care to mitigate face threat, is diminished. The difference between what is said and what is reported as said thus provides a potentially useful analytic window on the specific ways people use language to produce these different pragmatic effects. Here we focus in particular on what this contrast can tell us about the way people formulate and report on their agreements and disagreements with others, and what this can reveal about stance.

Reported speech could also be used as a way to represent people's stance on the specific matters being addressed in a conversation. Even if an utterance is not formulated as being an agreement or disagreement it can be clear from the context that it contrasts in some way with something somebody else said. This can be

achieved using turn-initial markers such as 'well' that can signal a speaker is going to produce something that is not well aligned with a previous turn or, as illustrated in Example 1, by using negation. If reported speech is used for the expression of opinions and disputes then markers of contrastive stance or position should be more common than in direct speech. This is of particular interest because these are devices that enable the current addressee to re-construct a version of the original prior turn that reported speech represents a response to.

## 3.2 Hypotheses

Three general hypotheses are identified for reported speech:

**1. Politeness:** The general politeness hypothesis is that people avoid the face-threat involved in direct disagreement with an addressee. Unless a current addressee is aligned in some way with the person(s) whose speech is being reported then the pressure for mitigation of negative responses is removed.[2] The general *politeness hypothesis* thus predicts that reported speech should tend to contain more exposed disagreement than direct speech.

**2. Self-Presentation:** Even where people are not disagreeing directly with their current addressee they might still wish to demonstrate that they understand that disagreement is a sensitive matter e.g., to avoid the inference that they are rude or combative. If people are sensitive to this then, all things being equal, they should not produce any more explicit disagreements in reported speech than they do in direct speech. Moreover, concerns about self-presentation should by definition affect 'self' more strongly than 'other' therefore we would expect fewer explicit markers of disagreement in self-reported than other-reported speech.

**3. Contrastive Stance:** A third general hypothesis is that people's primary concern when reporting on a prior conversation is to highlight the substantive differences between their own stance and that of others. The intuition here is that like ordinary utterances reported speech should ideally be newsworthy in some way (Goodwin, 1979); either to the current addressee as a means of highlighting a significant stance previously taken by the speaker, or to convey the newsworthiness of the reported speech to the people actually in the prior conversation. This leads to the prediction that reported speech should contain more explicit markers of stance or emphasis than direct speech; for example, by using turn-initial discourse markers such as 'well' or negations (Scott, 2002) as illustrated above in Example 1.

[2]Of course, it is possible that the current addressee might also take issue with the opinion or stance identified in the reported speech but this would become an issue for their *subsequent* response to the report not the format of the report itself.

## 3.3 Linguistic Features

In order to make quantitative tests of these predictions we now consider some potential indices of the different ways people can position direct and reported speech. In particular, discourse markers of (dis)agreement, stance, emphasis and contrast are considered. As explored in section 2.6.2, discourse markers are devices which can enable speakers to explicitly show the relationship between two utterances, such as highlighting that what is said is positioned in contrast to what has come before (Fox Tree and Schrock, 1999, 2002; Fraser, 1996; Heritage, 1984; Smith, 1998).

### 3.3.1 Markers of Agreement and Disagreement

The simplest case for analysis is where people explicitly position their turns as agreement or disagreement. This can be done with phrases such as 'You're wrong', 'I disagree', 'I don't agree' and 'You're right' or 'I agree'. Unfortunately, for the reasons outlined above these exposed forms, especially those associated with disagreement, are likely to be rare.

A second set of more indirect indicators are provided by cue words or discourse markers that are associated with agreement and disagreement but don't explicitly formulate a turn as such, taken from Walker et al. (2012c) (see section 2.6.1 for a more detailed discussion of how these markers were obtained). For our analysis then, more implicit markers of disagreement are: 'really', 'no', 'actually', 'but', 'so', and 'you mean'; and markers of agreement: markers of agreement: 'yes', 'I know', 'I believe', 'I think', and 'just' (Walker et al., 2012c).

A potential limitation of these indirect markers for the purpose of this study is that they are drawn from analysis of online discussion forums which are less dialogical than face-to-face interaction and where people may also tend to actively seek out disputes. It is also worth noting that, for example, the frequency of turn-initial 'yes' is not an unambiguous indicator of agreement; disagreement is often preceded by techniques including agrees (e.g. 'yes, but...'), delays and prefaces, such as, 'well' and 'hmm' (Kotthoff, 1993; Pomerantz, 1984a; Sacks and Jefferson, 1995). Clift (2006a) observes that 'well' can act as a buffer. Nonetheless, we assume that the relative distribution of these markers across different samples is indicative of the overall patterns of agreement and disagreement within them.

### 3.3.2 Update Markers

In addition to marking the fact of agreement and disagreement there are more subtle pragmatic markers that can signal an individual's knowledge state or stance with respect to the current conversational context. Here we use 'well' and 'oh', which we

gloss as *update markers* both of which are associated with signalling some form of contrast or sequential discontinuity in dialogue.

A turn-initial 'well' typically (but not exclusively) indicates that what follows will be in some way unexpected, unwelcome, discontinuous or contrary to a prior statement (Heritage and Clayman, 2010; Pomerantz, 1984b; Schegloff and Lerner, 2009; Schiffrin, 1988). As such it can signal that a forthcoming utterance is contrasting, unexpected or perhaps unwanted in substance, and which will lead to an update of the knowledge status.

A turn-initial *oh*, by contrast, typically (but not exclusively) acts as a reactive *change-of-state* token that indexes a responsive shift to a prior utterance through an update in the speaker's knowledge or awareness (Heritage, 1984, 1998). Schiffrin (1988) observes that *oh* often marks a shift in speaker orientation or stance, indicating a speaker's realisation that the hearer is not similarly aligned or oriented towards a proposition and may signal a potentially argumentative stance. Heritage also notes that 'oh' indexes a shift, identifying it as a 'change-of- state' token, highlighting an update in the speaker's knowledge or awareness (Heritage, 1984, 1998; Heritage and Raymond, 2005), which is supported by (Fox Tree and Schrock, 1999) who observes that it can signal a disjuncture, aiding hearers to integrate discourse.

### 3.3.3 Contrast, Emphasis and Expletives

Finally, in order to index the way in which the content of a turn is formulated or positioned with respect to another turn, we track negations ('not' and 'n't') and mid-turn contrastive conjuncts ('but' and 'though') as markers of contrast. The role of negation as a key phenomenon in relation to opinion and disagreement has been noted in the literature (Benamara et al., 2012; Scott, 2002) and is of particular interest here because of its use for the denial or rejection of statements; consequently, its role in rejection and disagreement, together with its inherent connection to the expression of alternatives or contrast, led to the inclusion of negation for our analysis. Adverbial emphasisers, such as 'really', 'surely' and 'clearly', are included as indicators of emphasis (Quirk and Crystal, 1985). The role of adverbial emphasisers as possible indices of disagreement (Scott, 2002) and for the expression of stance (e.g. conveying attitudes towards the content of a sentence), have been highlighted in the literature (Biber and Finegan, 1989; Conrad and Biber, 2000). We also track frequencies of a manually compiled list of common swearwords informed by previous studies and frequency data that surfaced from a sub-sample of the BNC dataset ('bastard', 'bitch', 'bloody', 'bollocks', 'fuck', 'piss off', 'shit' and 'wanker') which can be used for the expression of emotions, especially frustration, anger and surprise (Jay and Janschewitz, 2008).

## 3.4 Predictions

Building on the three general hypotheses presented above and the discussion of different markers of agreement, disagreement and stance we can summarise eight basic predictions:

1. Politeness: Markers of agreement should always be more common than markers of disagreement in all speech.
2. Politeness: Markers of disagreement should be more common in reported speech than direct speech.
3. Politeness: Expletives should be more common in reported speech than direct speech.
4. Self-Presentation: Markers of disagreement should not be more common in self-reported speech than direct speech.
5. Self-Presentation: Markers of disagreement should be less common in self-reported speech than other-reported speech.
6. Self-Presentation: Expletives should be less common in self-reported speech than other-reported speech.
7. Contrastive Stance: Update markers should be more common in reported speech than in direct speech.
8. Contrastive Stance: Contrast and Emphasis should be more common in reported speech than in direct speech.

## 3.5 Method

The corpus analysis used the spoken dialogue component of the British National Corpus (BNC), comprising approximately 10 million words. This sizeable collection of naturally occurring conversations offers scope to explore patterns of reported speech across a large sample. The transcripts include annotations for some key paralinguistic features such as laughing, overlapping speech and significant pauses, although the transcription conventions vary. Our analysis is based on the BNC's *s-units* which are sentence-like divisions of the transcribed utterances. We used SCoRE, a web interface for dialogue corpora, to gather our data from the BNC (Purver, 2001). It can be used to search for any regular expression, and for word or phrase repetitions, including repeats across sentence/turn boundaries.

For each set of markers their frequency in the BNC was gathered and analysed. Reported speech can be introduced in a number of ways, for example, 'I went', 'I says', 'he goes', 'she was like'. We focused on 'pronoun + said + report' as this produced a good sized dataset. Using the ScoRE interface (Purver, 2001) it was possible to extract all instances of 'I said' (5315 turns), 'he said' (3310 turns) and 'she said' (2579 turns), which were then checked by hand to ensure they were consistent samples of reported speech. A further 5315 turns were randomly selected from the spoken dialogue section of the BNC to provide a comparable sample of general direct speech.

In order to control for the possibility that reported speech tends to occur in particular dialogue contexts or with particular audiences (e.g., story-telling to friends) a second sub-sample of 500 turns of direct speech was selected from the same context by identifying the nearest preceding turn to an identified instance of self-reported speech ('I said') by the same speaker, that did not contain an instance of reported speech. This is referred to below as the *Local Context* sample.

The samples were analysed for a number of turn-initial features: agreement and disagreement markers, update markers 'oh' and 'well'. Turn-initial in the reported speech samples constituted what immediately followed I/(s)he said, while in the direct speech sample it was simply the initial words of the turns. Non-turn-initial features were also investigated: adverbial emphasisers (often indicators of stance or opinion markers), 'oh' (change-of-state tokens), negations and swearwords.

## 3.6 Results

In this section the results of the corpus analysis are presented.

### 3.6.1 Exposed Disagreement

As Table 3.3 shows, both exposed agreement and disagreement are rare, although exposed agreement is, as expected, more common than disagreement. Only 0.8% of the turns sampled contain strong expressions of disagreement whereas 5.2% contain strong expressions of agreement. Strikingly, over 97% of these instances of exposed agreement/disagreement occur in direct speech. This observation is clearly counter to the initial politeness hypothesis for reported speech and incompatible with the self-presentation hypothesis.

Chi Square analysis of the frequency of strongly exposed agreement and disagreement indicates that their distributions are different in reported and direct speech

| Phrase | Reported Speech | Direct Speech | Total |
|---|---|---|---|
| You're wrong | 6 | 17 | 23 |
| I disagree | 0 | 15 | 15 |
| I don't agree | 2 | 46 | 48 |
| You're right | 5 | 224 | 229 |
| I agree | 5 | 318 | 323 |

Table 3.3 Instances of Exposed Agreement and Disagreement in the BNC. *RS* = Reported Speech and *DS* = Direct Speech

($\chi^2_{(1)}$ = 15.23, p<0.01). [3] There is approximately a 7:1 bias toward overt expression of agreement over disagreement in direct speech compared with approximately 1:1 in reported speech. This suggests that although explicit, exposed disagreement is much less common in reported speech there is no particular bias in that context toward overtly positioning a relayed turn as agreement or disagreement.

### 3.6.2 Agreement and Disagreement markers

The distribution of turn-initial markers of agreement and disagreement identified by Walker et al. (2012c) for each subsample are shown in Tables 3.4 and 3.5.

| Marker | Direct Speech | (s)he said | I said | Context |
|---|---|---|---|---|
| Really | 3 | 5 | 4 | 1 |
| No | 173 | 128 | 190 | 12 |
| Actually | 3 | 5 | 2 | 1 |
| But | 85 | 63 | 51 | 13 |
| So | 108 | 30 | 17 | 19 |
| You mean | 0 | 0 | 0 | 0 |
| **Total** | **372** | **231** | **264** | **46** |
| Total turns | 5315 | 5889 | 5315 | 500 |
| % total turns | 7.00 | 3.92 | 5.00 | 9.20 |

Table 3.4 Frequency of Disagreement Markers

As Table 3.4 suggests, the overall frequency of markers of disagreement is higher in direct speech than all reported speech ($\chi^2_{(1)}$ = 48.3, p<0.01) and also higher in the

[3] Throughout we use p<0.05 as our criterion level but report computed probabilities to two decimal places for completeness.

| Marker | Direct Speech | (s)he said | I said | Context |
|---|---|---|---|---|
| Yeah/Yes | 647 | 139 | 181 | 26 |
| I know | 12 | 16 | 22 | 4 |
| I believe | 0 | 1 | 1 | 1 |
| I think | 31 | 22 | 27 | 3 |
| I just | 4 | 10 | 6 | 2 |
| **Total** | **694** | **188** | **237** | **36** |
| Total turns | 5315 | 5889 | 5315 | 500 |
| % total turns | 13.06 | 3.19 | 4.46 | 7.20 |

Table 3.5 Frequency of Agreement Markers

Local Context sample (i.e. preceding direct speech turn by the same speaker) than in the self-reported speech of the same speaker ($\chi^2_{(1)}$ = 16.22, p<0.01). Comparison of self-reported speech with other-reported speech (he/she said) shows markers of disagreement are less common in other-reported speech ($\chi^2_{(1)}$ = 7.22, p=0.01). These patterns are opposite to the predicted pattern for the Politeness and Self-Presentation hypotheses for reported speech.

The same pattern is observed for the markers of agreement. They are more common in direct than reported speech ($\chi^2_{(1)}$ = 489, p<0.01) and more common in the Local Context sample from the same speaker than in self-reported speech ($\chi^2_{(1)}$ = 7.63, p=0.01). They are also more common in self-reported speech than other-reported speech ($\chi^2_{(1)}$ = 12.2, p<0.01).

Overall the results show that explicit and implicit markers of agreement and disagreement are more common in direct speech than reported speech and more common in self-reported than other-reported speech.

### 3.6.3 Turn-Initial Update markers

The raw frequencies for the distribution of turn-initial update markers are provided in Table 3.6. The 'reactive' change of state token 'oh' is more common in reported speech than all direct speech ($\chi^2_{(1)}$ = 16.7, p<0.01) but there is no difference in frequency between self-reported speech and the Local Context turns by the same speaker ($\chi^2_{(1)}$ = 0.58, p=0.45). 'Oh' is however, slightly more frequent in other-reported speech (he/she) than self-reported speech ($\chi^2_{(1)}$ = 4.72, p=0.03).

As Table 3.6 shows, differences in the use of the 'prospective' update marker 'well' are more marked. It is approximately twice as common in reported speech as direct speech ($\chi^2_{(1)}$ = 70.9, p<0.01). Most of this difference is accounted for by

| Marker | Direct Speech | (s)he said | I said | Context |
|---|---|---|---|---|
| Oh | 170 | 292 | 218 | 17 |
| Well | 202 | 299 | 502 | 22 |
| **Total** | **372** | **591** | **720** | **39** |
| Total turns | 5315 | 5889 | 5315 | 500 |
| % total turns | 7.00 | 10.04 | 13.55 | 7.8 |

Table 3.6 Frequency of Update Markers

the use of 'well' in self-reported speech where it is approximately twice as common as in the Local Context speech turn by the same speaker ($\chi^2_{(1)}$ = 14.2, p<0.01) and approximately twice as common in self-reported speech than direct speech ($\chi^2_{(1)}$ = 80.3, p<0.01).

Overall, in contrast to markers of (dis)agreement, signals of updates are more common in reported speech. The use of the reactive 'oh' is more strongly associated with other-reported speech whereas the use of the prospective 'well' is associated with self-reported speech.

### 3.6.4 Contrast and Emphasis

The counts for markers of contrast and emphasis i.e. negations, contrastive conjunctives (but, though), adverbial emphasisers (actually, certainly, clearly, definitely, indeed, obviously, plainly, really, surely, for certain, for sure, of course) and common swearwords are provided in Table 3.7. For all these markers occurrences at any position within a turn were included for analysis.

| Feature | Direct Speech | (s)he said | I said | Context |
|---|---|---|---|---|
| Negation | 624 | 1300 | 1211 | 148 |
| Swearwords | 6 | 90 | 132 | 3 |
| Contrastives | 298 | 316 | 411 | 62 |
| Adverbials | 187 | 162 | 158 | 40 |
| **Total** | **1115** | **1868** | **1912** | **253** |
| Total turns | 5315 | 5889 | 5315 | 500 |
| % total turns | 20.98 | 31.72 | 35.97 | 50.6 |

Table 3.7 Frequency of Negations and Adverbial emphasises

It is immediately clear from Table 3.7 that swearwords are much more common in reported speech than in direct speech ($\chi^2_{(1)}$ = 92.5, p<0.01); they are also more common in self-reported speech than other-reported speech ($\chi^2_{(1)}$ = 76.8, p<0.01). Swearwords are also four times more common in self-reported speech than in the Local Context turns by the speaker ($\chi^2_{(1)}$ = 7.15, p<0.01).

Negations follow a similar pattern. They are approximately twice as common in reported speech as direct speech ($\chi^2_{(1)}$ = 266, p<0.01) and approximately twice as common in self-reported speech as other-reported speech ($\chi^2_{(1)}$ = 350, p<0.01). However, negations are less frequent in self-reported speech than in the Local Context turns by the same speaker.

Contrastive conjunctives are also more common in reported speech than direct speech ($\chi^2_{(1)}$ = 4.82, p=0.03) and more than twice as common in self-reported speech than in other-reported speech ($\chi^2_{(1)}$ = 25.79, p<0.01). However, like negations they are less frequent in self-reported speech than in the Local Context turns by the same speaker ($\chi^2_{(1)}$ = 13.3, p<0.01).

The pattern for adverbial emphasisers is different to the other markers of contrast. Emphasis is both slightly more common in direct speech than reported speech ($\chi^2_{(1)}$ = 5.31, p=0.02) and equally frequent in self-reported and other-reported speech ($\chi^2_{(1)}$ = 0.48, p=0.48). It is also approximately twice as common in the Local Context sample of the speaker (context sample) than in their self-reported speech. Overall, emphasis is slightly more common in direct speech overall and particularly common in turns introducing reported speech.

## 3.7 Discussion

Direct forms of disagreement occurred very rarely in the BNC. Thus it seems that directly stating that you disagree with someone is not common practice. Although the results show a clear preference for agreement over disagreement in direct speech, they also show that, contrary to the predictions of the politeness hypothesis, reported speech does not appear to be a context in which explicit disagreements are more likely to be exposed. On the contrary, people are far less likely to include explicit markers of agreement or disagreement in reported speech than in direct speech. Moreover, where they do formulate a reported utterance with an explicit marker it is equally likely to be agreement or disagreement.

Explicit makers of agreement and disagreement are rare of course and not an essential part of actually enacting an agreement or disagreement. However, the results show the same pattern for the less direct markers of agreement and disagreement identified by Walker et al. (2012c). Again, markers of both disagreement and

agreement are more common in direct speech than reported speech. Overall, it appears that reported speech is not a context in which disagreements are normally re-presented or rehearsed as disagreements.

These results also run counter to the hypothesis that the format of reported speech turns is constrained by concerns with self-presentation. The results are contrary to predictions 5,6 and 7. Although the self-presentation hypothesis predicts that disagreement should not be more common in reported speech, it is incompatible with the observation that it is more common in direct speech and more specifically more common in self-reported speech than other-reported speech. A self-presentation account is also difficult to reconcile with the observation that ostensibly taboo swearwords are more common in direct than reported speech; self or other.

The hypothesis that provides the best fit to the preceding results is Contrastive Stance. The results suggest that reported speech is not used for the re-presentation of (dis)agreements, or at least not in the same way in which they are actually enacted in direct speech. Firstly, the update markers 'Oh' and 'Well' appear to be quite strongly associated with reported speech. This suggests people are deliberately highlighting moments of change more than they actually mark them in direct speech. Although not directly predicted the additional observation that people are more likely to 'well'-preface a self-report of their own remarks and 'oh'-preface reports of another's remark suggests individuals position themselves as delivering updates and report on others receiving them. This asymmetric highlighting of changes in epistemic stance fits with a concern to re-present the newsworthy and contrastive elements of prior conversations. Within these reports what is selected for inclusion also appears to focus on the substance of a dispute, i.e. on expressions of contrast and features that indicate shifts in stance. This is compatible with the relatively low frequency with which 'meta' agreement and disagreement markers are used. It is also compatible with the increased use of use of negations and contrastive conjunctives.

However, there are also some challenges to the Contrastive Stance hypothesis in the data presented above. It doesn't directly account for the observation that swearwords will be used more frequently unless these are also construed primarily as markers of contrast, perhaps acting as and emphasis device. This is plausible but post-hoc. Also, its prediction that markers of emphasis should be more common in reported speech is not borne out. The results show that the turn preceding reported speech (the 'Local Context' turn) does tend to include emphasis so this might reflect a marking of stance but again, this is a post-hoc explanation. It appears that highlighting points of contrast and representing stance and shifts in assessed parameters are key functions of reported speech. This aligns with Clift (2006a), in that it suggests that there is an evidential interactive function being enacted through the use of reported speech.

While this study shows that reported speech is not used to re-present how disagreements were enacted, it is possible that other forms of report may. The dataset we worked with predominantly included direct reported speech or quotatives ('he said cats are bad'), but also some indirect reported speech ('he said that cats are bad'). Further work to investigate how the more descriptive indirect reports, and the wider gamut of reported thoughts might be used to re-present disagreement may provide further insights into the reporting of disagreement.

## 3.8 Implications

In this chapter a corpus study of reported and direct speech within the BNC was examined for a variety of disagreement and stance markers. In line with literature, it was found that exposed disagreement is rare in conversation. Even within the context of reported speech, in which the potential face threat of an utterance is diminished, directly challenging a speaker, or overtly disagreeing with them is avoided. Consequently understanding the effect of exposed agreement and disagreement may be difficult to assess from corpus approaches. Although socially problematic, the potential benefits of disagreement, particularly for substantive discussion are wide reported. This leads to the question, what role could exposed disagreement play in debate contexts? Does, or can, disagreement help to advance the dialogue and lead to a more deliberative exchange, or must it always be socially problematic? What is constructive engagement in discussion contexts and what is the role of disagreement in this process?

The corpus of reported speech included higher frequencies of features which can emphasises speaker position, such as contrast and negation. Thus rather than doing disagreement in direct ways, perhaps making explicit the stance - i.e. the difference between speaker positions - is a more important feature to examine when trying to understand how disagreement is enacted in dialogue.

# Chapter 4

# Shifting Opinions: an Experimental Approach to Disagreement

## 4.1 Introduction

Disagreement is understood to be socially problematic; it also rarely surfaces in naturally occurring conversation. In this chapter an experiment was designed to facilitate the direct manipulation of the occurrence of exposed (dis)agreement, also referred to as impolite disagreement elsewhere in the literature, and to rack the effects on the subsequent dialogue. This is the first experiment to directly manipulate the occurrence of exposed agreement and disagreement in dialogue.

Previous studies on disagreement take a distributional or corpus based approach at evidencing and analysing instances of disagreement in interaction (Abbott et al., 2011; Holtgraves, 1997; Misra and Walker, 2013; Walker et al., 2012c). These studies have provided valuable insights into the ways in which these complex social interactions are handled in different contexts, and given rise to various theories on how we process, respond to and mitigate the impact of disagreement. However, the literature also highlights that exposed disagreement rarely surfaces in naturally occurring conversation, as confirmed by the corpus study presented in chapter 3.

This chapter outlines an experimental approach for investigating disagreement, which provides the opportunity to manipulate the occurrence of *exposed* (dis)agreement in dialogue. By *exposed*, we refer particularly to direct and unequivocal presentations of agreement and disagreement, such as 'I agree' and 'You're wrong'. However, we also explore less direct markers, which can, but do not always function in a (dis)agreement capacity. For example, turn-initial 'yes' and 'no', can and are often used to signal agreement and disagreement, however, the function of these markers is context specific and dependent on the preceding content (for example a 'no' following a negative statement can function as agreement). In more general terms,

disagreement, for the purposes of this study, refers to the presentation of a stance position that is incongruous, misaligned or somehow in opposition to or 'at odds' with what has come before. Conversely, an agreement is characterised by an utterance that serves to reinforce a prior contribution, denote alignment or concordance and acceptance of a prior contribution.

### 4.1.1 Motivation for the study

Expressing a view in opposition to that of your interlocutor, i.e. one that contradicts, challenges, refutes or denies their position, attitude or contribution, can be socially problematic. Direct challenges to a speaker or disagreeing with their assertion in dialogue can constitute, in Brown and Levinson's terminology, what is known as a *Face Threatening Act*, that is to say it can threaten the hearer's public self-image. Consequently, as demonstrated in chapter 3, disagreement, especially when done in a direct manner, is rare in conversation. This means that it is difficult to assess what affect it has upon a dialogue. An experimental approach has the advantage that it allows us to directly manipulate the occurrence of exposed (dis)agreement and track it's effects on the subsequent dialogue.

As discussed in section 2.2.2 disagreement can signal social intimacy, lead to novel contributions in problem solving talks, and be essential in developing the deliberative quality of a dialogue. So although disagreement, particularly when executed impolitely, tends to be problematic, for certain contexts, such as problem solving and discussion tasks, it can be essential in advancing the deliberative quality of a dialogue. Chiu (2008) also suggests that agreement can be potentially detrimental to a dialogue, but the problematic aspects of agreement are not well reported in the literature; this gives rise to a secondary question, 'what affect does exposed *agreement* have upon a dialogue?' If it is problematic, how and in what ways does this manifest?

If disagreement encourages novel contributions does agreement, conversely, stifle them? If people are too readily agreeing, does this prevent more involved discussion that could lead to shifts in stance or the development of new contributions? In order to understand the effects of both exposed agreement and disagreement, an experiment was designed that enabled the manipulation of such features under controlled conditions. This is the first experiment to directly manipulate the occurrence of exposed agreement and disagreement in dialogue.

The causal *effects* of exposed disagreement on the subsequent trajectory of a dialogue has not previously been directly tested. One key reason for this is the practical difficulty of manipulating linguistic features in a live dialogue. Here a technique introduced by Healey et al. (2003) is used that takes advantage of the

potential of text-chat for enabling selective manipulation of people's turns without their awareness.

### 4.1.2 Politeness and Accommodation Theory

One argument for the scarcity of disagreement in dialogue is anchored to the concept of politeness. Politeness Theory builds upon Ervin Goffman's concept of *face*. Goffman (1967a) defines face as 'the positive social value a person effectively claims for himself' through interaction and offers a model of co-operation that is enacted when an individual's face or social value is threatened during interaction. Goffman stresses the co-operative nature of facework: 'When a face has been threatened [...] lack of effort on the part of one person induces compensative effort from others' (Goffman, 1967a). This mutual co-operation and shared consideration in interaction has also been located as a central notion for Politeness Theorists (Brown and Levinson, 1987; Watts, 2003).

Politeness Theory suggests that interlocutors minimise disagreement to save face, employing strategic conflict avoidance techniques to mitigate the effect of any disagreement that may surface (Leech, 1980). However, Accommodation Theory would posit that if someone is agreeable their conversational partner would match them in this convivial approach, whereas if they are adopting a discursive or even combative linguistic style, then their conversational partner would be likely to adopt a similar tact and synchronicity would become more exaggerated (Giles and Smith, 1979). Accommodation Theory posits that interlocutors adopt strategies of *convergence* to integrate and identify socially with another (Giles et al., 1991); this involves the adoption of linguistic similarities and leads to perceived communicative effectiveness (Giles and Smith, 1979) and cooperativeness (Feldman, 1968). Conversely, speech *divergence* reflects distancing from the co-conversant and can surface when confronted with perceived differences to the co-conversant.

#### Disagreement and Constructive Engagement

As demonstrated in section 2.6.1, although disagreement has often been regarded as socially problematic, a number of potential benefits have also been countered. In particular contexts such as task orientated dialogue, problem solving talk and work place interactions, disagreement is identified as an essential process that can lead to greater clarity (Colman et al., 2011), novel contributions Chiu (2008), and is both acceptable and expected (Angouri, 2012; Angouri and Locher, 2012). However, it is found that direct disagreement is still very rare in these contexts, with more polite formulations being preferred (Angouri, 2012; Chiu, 2008).

Conversely, researchers suggest that agreement, rather than disagreement, can pose problems to the success of a dialogue (Chiu, 2008). Martin and White (2007) argue that agreement can serve a contractive function, serving to close down the possibility of further discussion. Issuing only agreements can lead to a lack of mutual intelligibility and prevent further elaboration on the position an interlocutor is presenting and exposition of the reasoning underpinning it. Intuitively this would suggest that agreement can serve to close down a dialogue; does it therefore follow that disagreement could, conversely, serve to open up the dialogic space? If disagreement encourages novel contributions does agreement, conversely, stifle them? If people are too readily agreeing, does this prevent more involved discussion that could lead to shifts in stance or the development of new contributions?

Consequently, disagreement ought to be a catalyst or precursor to a potential shift in position, as it signals a direct challenge to a held idea, which in turn may be retained, re-negotiated or more fundamentally re-conceived. This, together with the findings by Chiu (2008), suggests that disagreement can play an important role in the deliberation and problem solving process. However, the exact effect of direct and unmitigated disagreement has been little explored, owing to its highly marked nature and therefore rare occurrence. Furthermore, a comparison of the exact interactional effects of agreement and disagreement have hitherto not been fully explored. In this chapter, and experimental approach which facilitates such a direct comparison is presented.

### 4.1.3 Predictions

Given the literature we would expect that exposed disagreement would be especially problematic; it should instigate additional work being done in the interaction and more instances of repair. Insertions of exposed disagreement should be more disruptive than exposed agreement insertions, which should in turn facilitate more agreement. Assuming speakers are being co-operative, all things being equal, then disagreement should lead to more hedging and mitigation in order to manage the disagreement and minimise face-threat. However, it may also lead to additional stance shifts, or the consideration of more alternatives during the discussion dialogues.

## 4.2 Agreement and Disagreement Fragment Experiment

In order to assess the impact of exposed (dis)agreement, an experiment was designed in which instances of exposed (dis)agreement were artificially inserted into a dialogue.

Turn-initial discourse markers such as 'No', 'But' 'you're wrong' and, 'I disagree' can highlight instances of disagreement within a conversation. Similarly, 'Yes', 'And', 'I agree' and 'you're right' can serve as indicators of agreement, or reinforce congruence. These eight fragments were selected because they provide a range of exposed, direct (dis)agreement and more subtle markers that *can* be used in (dis)agreement. Using the DiET chat tool it was possible to alter a participant's turn before it was relayed to their conversational partner. This means that the conversational partner would receive the updated turn with the intervention fragment inserted at the beginning of the original turn, while the message sender would only see their original turn.

| Sender (A) chat log | Receiver (B) chat log |
|---|---|
| A: What do you think? | A:What do you think? |
| B: Tom should go | B: Tom should go |
| A: Tom's the pilot!!! | A But Tom's the pilot!!! |

Table 4.1 Example of participants' respective views of intervention turns

### 4.2.1 Hypotheses

1. Accommodation Theory: The general accommodation hypothesis is that dialogue partners match linguistic and discursive style. Thus the general *accommodation hypothesis* predicts that the insertion of agreement fragments will elicit additional instances of agreement, while the insertion of disagreement fragments will elicit additional instances of disagreement.

2. Politeness: The general politeness hypothesis is that face-threatening acts are socially problematic and should result in compensatory action being taken to redress and mitigate the situation. The general *politeness hypothesis* thus predicts that inserting disagreement fragments into a dialogue should lead to more work being done and more co-operation and consideration being displayed; this may result in increased effort when formulating responses (higher number of self-edits) and more clarification requests, expressions of agreement and other routinised polite sequences.

3. Constructive Disagreements: The general constructive disagreement hypothesis is that disagreement is essential for advancing the deliberative quality and problem solving aspects of dialogue. The *constructive disagreement hypothesis* thus predicts that people will respond constructively to disagreement. The specific predictions for particular response measures are a much lower level issue, but we would expect the insertion of disagreement fragments to lead to

increased deliberation taking place which lead to a higher number of shifts in stance and alternative solutions considered over the course of a dialogue.

### 4.2.2 Method

Pairs of participants were seated at separate computers in adjacent rooms and given an instruction sheet to read detailing the balloon task. Participants are presented with a fictional scenario in which an hot air balloon is losing altitude and about to crash. The only way for any of three passengers to survive is for one of them to jump to a certain death. The three passengers are: Dr. Nick Riviera, a cancer scientist, Mrs. Susie Derkins, a pregnant primary school teacher, and Mr. Tom Derkins, the balloon pilot and Susie's husband. Participants are told to take as much time as they need to read the summary of the situation and then discuss with their partners via a chat tool set up on the computer at which they are seated, and attempt to come to a conclusion over who should jump from the balloon. The advantages of this task are that it is effective at generating debates between subjects and involves articulations of agreement and disagreement as they attempt to come to a conclusion. There is also plenty of scope for deliberation and shifts in stance.

### 4.2.3 Participants

Seventy-two participants were recruited, 46 female and 26 male, with the majority being undergraduate and postgraduate students at the University of London. Participants were invited to attend with someone who they already knew. They were recruited in pairs to ensure that inter-pair participants were acquainted. For a couple of experiments if one participant didn't show up a stand in was recruited last minute, and in these exceptions, which are marked in the data, the pair were not previously acquainted with each other. Each participant was paid at a rate of £7.50 per hour for participating in the experiment, or if they were a Psychology student at Queen Mary University of London then they could receive course credits in lieu of payment.

### 4.2.4 Materials

The participants communicate via a specially programmed chat tool, similar to other instant messenger interfaces they may have used previously. The Dialogue Experimental Tool kit (DiET) chat tool is a text-based chat interface facilitating real time manipulations of the dialogue. It is possible to programme several different types of interventions using the chat tool: turns may be altered prior to transmission, turns may not be relayed, and additional turns may be added, (e.g. Healey et al. (2003), insertion of spoof clarification requests).

These manipulations occur as the dialogue progresses, thus making them minimally disruptive to the sequence of dialogue. The DiET chat tool is built in Java and consists of a server console and user interface. Participants are faced with a text box displaying the conversation history and a smaller text box into which they can type. Participants can type simultaneously and their message is relayed to their conversation partner by use of the ENTER key. The server time stamps and stores all key presses. All turns are passed to the server before being transmitted to the other participant, thus making it an intermediary between what the participants type and what they receive. Turns can be automatically altered, removed or inserted by the server before they are relayed.

### 4.2.5 Design

The experiment is conducted in pairs; there were 12 dyads for each condition. Pairs of participants were presented with a discussion task and instructed to discuss for 30 minutes and attempt to come to an agreement. Each pair of participants was assigned to a condition at random. There were three experimental conditions. Please note, what we gloss here as the *Agreement* and *Disagreement* conditions, are named as such because the inserted fragments in each condition *can* index disagreement, however, it is acknowledged that the more indirect fragments do not consistently perform this function.

- **Control condition**: Participants are welcomed and briefed before being sat at their respective computers, which were situated in adjoining rooms. They receive their task instructions on a piece of paper and can start when they are ready. They are instructed to discuss the scenario and attempt to come to an agreement on who should jump from the balloon for 30 minutes. No interventions are performed by the server; participants receive the dialogue turns exactly as they were typed.

- **Agreement condition**: Initial procedure is exactly the same as the control condition. Participants receive the dialogue turns exactly as they were typed, except for every fourth turn when one of the following fragments is inserted position: ‘you’re right’, ‘I agree’, ‘yes’, ‘and’.

- **Disagreement condition**: Initial procedure is exactly the same as the control condition. Participants receive the dialogue turns exactly as they were typed, except for every fourth turn when one of the following fragments inserted at turn-initial position: ‘you’re wrong’, ‘I disagree’, ‘no’, ‘but’.

### Frequency of interventions

A small scale pilot study was conducted to assess how frequently the insertions could be made. It was important that enough interventions could be achieved over the course of a dialogue so that their effect could be reliably tested, but without causing participants to orientate themselves towards the intervention. Particularly due to the marked nature of the inserted fragments inserting an explicitly formulated disagreement such as 'you're wrong' too frequently became problematic. In an initial pilot experiment only the most explicit formulation case for each condition was used (i.e. 'I agree' and 'I disagree'). However, to avoid becoming obviously repetitive and due to the marked nature of insertions, it was not possible to insert the fragments regularly enough. This led to the decision to rotate four markers for each condition, two explicit formulations ('I agree'/'I disagree' and 'you're right'/'you're wrong') with two further less explicit markers of agreement or disagreement added to provide a more varied, and therefore more naturalistic intervention ('yes' and 'and' in the agreement condition and 'no' and 'but' in the disagreement condition). This combination of markers were piloted, with manipulations carried out every fourth turn issued by each speaker with no perceptible disruption to the dialogue nor orientation by the participants to the interventions. Triggering interventions every fourth turn enabled enough data to be gathered to test the experimental condition, whilst minimising the risk of making the intervention noticeable to participants. No intervention was made if the turn consisted of only one word, or the turn started with the same text as featured in the insertion fragments. This was to avoid the production of particularly nonsensical turns such as 'you're wrong I agree'. The fragments were cycled through in order but the exposed (dis)agreement fragments ('you're wrong/right', 'I (dis)agree') appeared half as often due to their marked nature.

## 4.3 Results

Data was gathered both directly from the chat tool which logged various features such as typing time, number of self-edits, i.e. use of the backspace and delete key and temporal data, as well as the transcripts themselves, which were analysed for linguistic features and frequencies. All interventions were removed before frequency counts were conducted to ensure that only markers and features actually contributed by the participant were counted. Additionally the resulting transcripts were hand coded for clarification requests and *stance shifts*, explained in more detail below.

### 4.3.1 A note on terminology

**Turn:** For the purpose of the experiments presented in this thesis, a turn constitutes the text relayed in a single message, meaning what is delineated by the ENTER key.

**Stance position:** For the purpose of the experiments presented in this thesis, the stance position refers to a participant's current, and publicly stated, opinion on which characters should be saved/killed in the balloon task scenario.

**Intervention Turn (IT):** The IT refers to the turn issued by a speaker which has had a Turn-initial intervention fragment inserted before the actual typed message.

**Intervention Reply Turn (IRT):** The IRT refers to the next turn issued by the speaker who receives the Intervention Turn. This is not always the next sequential turn after the IT, as the speaker whose turn contained the IT may issues another turn.

**Clarification Requests:** The transcripts were hand coded for Clarification Requests (CR), a form of repair in which speakers signal a need for further information, typically due to a lack of full comprehension of a previous utterance. This was done by a single annotator, blind, and all labelling indicating which condition a file belonged to was removed. CRs were hand labelled in the dataset, based on Purver et al. (2003) schema (included in Appendix section A.1), example provided in Table 4.2.

| | | | |
|---|---|---|---|
| Turn 1: | P1 | **you're wrong** or maybe we are just going by gender stereotypes.. the feminist in me is screaming | IT |
| Turn 2: | P1 | haha | |
| Turn 3: | P2 | what if thats the whole point | IRT |
| Turn 4: | P1 | sorry what if....? | CR |
| Turn 5: | P1 | susie jumped? | CR |

Table 4.2 Example of Reply Turn labelling

### 4.3.2 Coding for shifts in stance position

The transcripts were hand coded for shifts from one stance position to another regarding who to throw off of the balloon, i.e when a participant changed their point of view over who to sacrifice or save. There were seven potential stance states that cover all the possible combinations of who to save and who to sacrifice. The range of possible stances:

1. Undecided
2. Save Susie but undecided on who should die

3. Save Nick but undecided on who should die
4. Save Tom but but undecided on who should die
5. Sacrifice Susie (and therefore save the other two)
6. Sacrifice Nick (and therefore save the other two)
7. Sacrifice Tom (and therefore save the other two)

The annotation was done by a single annotator, blind, and all labelling indicating which condition a file belonged to was removed. A participant's stance position was carried over to the next turn, unless their current turn provided new information that contradicted the previous stance position, in which case the stance position was updated. By comparing the current stance position with the participant's last stance position it was possible to calculate whether or not a shift in stance position had occurred, the total number of alternative stance positions considered and a count of total stance position shifts.

A second annotator coded a sub-sample (approximately 10% of the total data) of the experiment data, three experiment transcripts taken from each condition, comprising 488 turns of dialogue. The annotators agreed on 87.70% of the codes, and intercoder reliability was calculated using the using Krippendorff's alpha (for nominal data) ($\alpha$= 0.70).

The annotation scheme was quite complicated and demanding, and required very minute changes in participant stance position to be tracked, which were not always clear from the text itself . As such, for future work a more simplistic coding scheme is recommended to achieve better intercoder reliability.

### 4.3.3 Overview of dataset

Table 4.3 displays the descriptive data for the turn, word and character counts for each condition.

| | Condition | | |
|---|---|---|---|
| Avg. | Control | Agreement | Disagreement |
| Turns by Dyad | 86.71 | 63.17 | 70.79 |
| Words by Dyad | 587.67 | 555.58 | 535.08 |
| Char. by Dyad | 2938 | 2797 | 2710 |
| Words per turn | 7.41 | 9.49 | 9.11 |

Table 4.3 Summary of average typed data per condition

Both intervention conditions result in fewer overall turns than the Control condition, but this was particularly the case, and statistically significant, with the Agreement condition (positive and agreement insertions, such as *yes* and *I agree*). Although the Agreement condition features fewer turns than the Control condition, there are more words per turn on average in the Agreement condition. A non-parametric Kruskal Wallis test confirms a significant overall effect of Condition on the turns typed in the dialogues ($H_{(2)}$ = 6.34, p<0.04).[1] Subsequent planned pairwise comparisons with the Dunn's test showed a significant increase in the number of turns per dyad in the Control condition compared to the Agreement condition (p<0.05). There is an overall effect of condition on the distribution of average words per turn, as confirmed by a non-parametric, Kruskal Wallis test ($H_{(2)}$ = 6.55, p<0.04). Subsequent planned pairwise comparisons with the Dunn's test showed a significant increase between Agreement and Control conditions (p<0.03).

### 4.3.4 Message construction

| Condition | Typing Time | Self-edits |
|---|---|---|
| Control | 11850 | 6.98 |
| Agree | 16210 | 6.97 |
| Disagree | 13484 | 7.51 |

Table 4.4 Table depicting mean Typing Time and number of Self-edits (delete key presses), per turn, per condition

Table 4.4 shows the average typing time in milliseconds and the number of self-edits per turn. Self-edits are represented by the number of times the delete key is pressed during turn construction. A non-parametric Kruskal Wallis test finds an omnibus effect of condition on the number of self-edits during turn construction ($H_{(2)}$ = 40.92, p<0.01), with planned pairwise comparison revealing significant difference between the Agreement and Disagreement conditions (p<0.01). An overall effect of condition on typing time is confirmed by a non-parametric Kruskal Wallis test ($H_{(2)}$ = 99.28, p<0.01), with planned pairwise comparison revealing significant difference between the Agreement and Control conditions (p<0.01).

### 4.3.5 Message content

The following tables highlight differences in the content of the dialogues, such as Clarification Requests and instances of exposed and potential disagreement.

[1] Throughout we use p<0.05 as our criterion level but report computed probabilities to two decimal places for completeness.

### Clarification Requests

| Condition | Total Number of CRs | Mean CRs per dyad |
|---|---|---|
| Control | 10 | 0.42 |
| Agreement | 13 | 0.54 |
| Disagreement | 50 | 2.08 |

Table 4.5 No. of Clarification requests by Condition

Table 4.5 shows the number of Clarification Requests by condition. The Disagreement condition has a significantly higher number of Clarification Requests than Control condition and Agreement condition. A non-parametric Kruskal Wallis test confirms an overall effect of Condition on the number of Clarification Requests in the dialogues ($H_{(2)}$ = 12.03, p<0.01). Planned pairwise comparison showed a significant increase between Control and Disagree conditions (p<0.01) and Agree and Disagree (p<0.02).

### Instances of exposed and potential (dis)agreement

Table 4.6 shows the frequencies of turn-initial exposed and potential (dis)agreement markers. The markers included here are the same ones that feature in the fragments that were artificially inserted during the experiment.

When considering all disagreement markers combined ('no', 'but','you're wrong', 'I disagree'), a non-parametric Kruskal Wallis test shows no reliable effect of condition on the frequency of disagreement markers in general ($H_{(2)}$ = 1.39, p=0.50). Similarly, a non-parametric Kruskal Wallis test shows no reliable effect of condition on the frequency of combined agreement markers ('yes', 'and','you're right', 'I agree') ($H_{(2)}$ = 2.71, p=0.26).

However, when exposed and indirect markers are pooled separately, the results are different. There is an omnibus effect of condition on the frequency of combined exposed markers of agreement and disagreement ('you're wrong', 'I disagree', 'you're right', 'I agree') according to a non-parametric Kruskal Wallis test ($H_{(2)}$ = 9.74, p<0.01).

When exposed agreement and disagreement markers are considered separately, a non-parametric Kruskal Wallis test shows a significant omnibus effect of condition on the frequency of turn-initial exposed markers of disagreement ('you're wrong', 'I disagree') ($H_{(2)}$ = 11.25, p<0.01). Subsequent planned pairwise comparisons with the Dunn's test showed a significant increase in the number of instances of exposed

| Turn-initial | Control condition | Agreement condition | Disagreement condition |
|---|---|---|---|
| Exposed (dis)agreement | | | |
| I agree | 2 | 5 | 10 |
| You're right | 0 | 0 | 0 |
| I disagree | 0 | 0 | 3 |
| You're wrong | 0 | 0 | 3 |
| Totals: Indirect (dis)agreement insertions | 2 | 5 | 16 |
| Yes | 170 | 124 | 139 |
| No | 29 | 23 | 35 |
| And | 103 | 51 | 55 |
| But | 119 | 81 | 77 |

Table 4.6 Table providing frequency data of turn-initial content of messages relayed during experiment dialogues.

| TI Update Marker | Control condition | Agreement condition | Disagreement condition |
|---|---|---|---|
| Well | 14 | 24 | 15 |
| Oh | 120 | 58 | 97 |
| **Total** | 134 | 82 | 112 |
| Total turns | 1893 | 1396 | 1506 |
| Percentage | 7.08% | 5.87% | 7.44% |

Table 4.7 Turn-Initial Update Markers

disagreement markers in the Disagreement condition compared to the Control condition (p=0.01) and the Agreement condition compared to the disagreement Condition (p=0.01).

The frequencies of turn initial 'and' and 'but', markers which can act as continuer and facilitate the structural continuation of contributions across turns are generally lower in the Control Condition than the Agreement or Control conditions, however, a non-parametric Kruskal Wallis test shows no reliable effect of condition on the frequencies of continuers ($H_{(2)}$ = 3.43, p=0.18).

#### Turn-Initial Update Markers: *Oh* and *Well*

Table 4.7 shows that Agreement condition featured fewer update markers proportionally than the Control condition, while the Disagreement condition featured slightly more. The frequency of turn-initial 'Oh' in particular is higher in Control condition, with Disagreement condition frequencies more closely resembling the Control than Agreement condition. However, a Kruskal Wallis test does not deem the difference a statistically significant effect of condition on the frequency of turn-initial *Oh* ($H_{(2)}$ = 2.94, p=0.23).

### 4.3.6 Intervention Reply Turns

In order to ascertain more clearly what effect the interventions had upon the dialogue, the following table displays descriptive statistics for the *reply turns*. These were counted as the next available turn by the participant who received the intervention. This was not always the next sequential turn, as for our experiment turns are demarcated by the use of the ENTER key.

#### Typing time

The average typing time of the intervention replies is significantly longer (mean value 24.46 seconds) than the typing time of non-intervention replies (mean value 15.59seconds). A General Linear Mixed Models analysis showed there was no interaction between condition and main effect, no main effect of condition ($F_{(1.76)}$ = 0.23, p=0.64), but a simple main effect of Intervention reply. There was no effect generally across conversations according to condition, the observed difference was specific to the intervention replies, not dependent on the nature of the intervention, i.e. meaning a disagreement or agreement insertion.

#### Time elapsed between speaker turns

We also captured how much time elapsed between when the enter key was hit by participant 1 relaying the *intervention turn*, and when the enter key was hit by participant 2 relaying the *Intervention Reply Turn*. It should be noted that this was not always the next consecutive turn, as participant one could send multiple turns before participant 2 responded.

Table 4.8 shows that average time elapsed before a reply is issued is significantly longer in the Intervention Reply Turns than in the control baseline - i.e. standard time elapsed between replies when no interventions are inserted ($F_{(2.57)}$ = 3.89, p=0.03).

| | Time Elapsed (in seconds) |
|---|---|
| Control condition (Baseline Reply) | 10.63 |
| Agreement condition (Intervention Reply) | 16.13 |
| Disagreement Condition (Intervention Reply) | 16.12 |

Table 4.8 Time elapsed between speaker turns

### 4.3.7 Stance shifts and Alternatives Considered

The experiment transcripts were also hand coded for *stance shifts*, i.e. when a participant voices a departure from one held opinion to an alternative regarding who should jump from the balloon.

| Condition | Total | Median | Mean | St. Dev. |
|---|---|---|---|---|
| Control | 175 | 7.5 | 8.33 | 3.96 |
| Agree | 248 | 11 | 10.33 | 4.88 |
| Disagree | 175 | 6 | 7.29 | 4.31 |

Table 4.9 Total number of stance position changes and averages per participant

The total number of stance position shifts and average per participant by condition are shown in Table 4.9. The median number of stance position changes per participant is significantly effected by condition ($\chi(2)$ = 6.91, p=0.03). A Median Test was conducted as the variance is not approximately equal across samples, being much larger for the agreement condition. This result suggests that the Disagreement condition tends to reduce the number of alternatives people will consider and the agreement condition tends to increase it.

There is no correlation between the length of the conversation (in turns) and the number of state changes (Kendals Tau = -0.007, p = 0.94), so the significance is not related to nor skewed by the fact that the Agreement condition contains longer dialogues, i.e. it is not just about how much participants talk.

In table 4.10, we provide summary data for the number of possible alternatives considered by a participant (out of a total of 7 stance states, i.e., "Kill Susie"). A non-parametric Kruskal Wallis test shows the difference in the number of possible alternatives considered was not significant affected by condition ($H_{(2)}$ = 3.62, p=0.16).

| Condition | Mean | Standard Deviation |
|---|---|---|
| Control | 4.65 | 1.18 |
| Agree | 5.30 | 1.26 |
| Disagree | 4.60 | 1.26 |

Table 4.10 Number of mean possible alternative solutions considered per participant by condition

## 4.4 Discussion

The general *accommodation hypothesis* predicted that the insertion of agreement fragments would elicit additional instances of agreement, while the insertion of disagreement fragments would elicit additional instances of disagreement. The turn-initial frequency data shows an increase in exposed agreement *and* impolite disagreement in the Disagreement condition when combined, and for increased exposed disagreement when considered separately, the Agreement condition does not contrast significantly with the Control in frequencies of exposed agreement *or* disagreement. Thus the general trend is counter to our Accommodation hypothesis, which anticipated that agreement would lead to more agreement while disagreement would engender more disagreement.

Although there is a notable absence of instances of exposed disagreement in the Agreement condition, with the comparative frequency in the Disagreement confirming a significant effect of condition, on closer inspection, what the so-called turn-initial exposed disagreement were doing was more complex. Rather than signalling an impolite disagreement, some of the instances counted were actually examples of quotatives, with participants repeating back the intervention content as a repair initiation. Therefore, a third of the instances of naturally occurring (i.e. not artificially inserted intervention fragments) turn-initial exposed disagreement markers in the Disagreement condition are actually instances of repair, rather than disagreement. To assess whether these repair instances had skewed the statistical tests, they were re-run with the falsely classified instances of exposed disagreement removed. A Kruskal Wallis test still found that there was a significant omnibus effect of condition on frequencies of exposed agreement and disagreement combined ($H_{(2)}$ = 7.10, p=0.029), with a pairwise comparison confirming a significant increase in the frequency of exposed (dis)agreement markers in the Disagreement condition compared to the Control condition (p=0.027). When the exposed disagreement marker frequencies considered separately there was also an omnibus effect of condition ($H_{(2)}$ = 6.35, p<0.05), however, a pairwise comparison using Dunn's test could not confirm a significant increase in exposed disagreement in the Disagreement

condition compared to the Agreement and Control conditions (Adjusted significance: $p<0.087$).

As shown in the following excerpt from an experiment transcript, the exposed disagreement is incongruent, jarring and provokes a repair sequence. The respondent quotes back the source of trouble, indicated by the asterisks in the example below, which were then falsely counted as naturally occurring turn-initial disagreement. The artificial insertions are shown in square brackets:

*Example 1*
A: Pros of keeping the doctor alive
A: [you're wrong] cures cancer
B: [no] you're wrong?*
A: What about?
B: no, I don't understand what you just said
B: You're wrong cures cancer?*
A: The doctor, if still alive will be about to discover the sure for the 'most common types of cancer'

This example demonstrates the disruptive nature of the inserted disagreement fragment; it disrupts the dialogue and is deemed incongruous enough for participant B to comment on, while participant A simply carries on with the conversation. This occurred several times in the dataset, however, only ever with the exposed disagreement fragments and never with the exposed agreement fragments. In line with the literature we found that exposed disagreement is especially problematic and on one occasion the insertion was so problematic that it was directly referenced and quoted by a participant, with both participants being alerted to the intervention.

*Example 2*
A: imagine how many scientists in the world
B: you're wrong theres a lot
A: i'm wrong?
B: what?
A: you said 'you're wrong theres a lot'
A: [no] what am i wrong about?

The Disagreement condition featured a significantly higher number of clarification requests. The Constructive Disagreement hypothesis anticipated that this would signal additional work being done by participants trying to more fully understand one

another's point of view. However, it is possible that the clarification requests are more clausal clarifications than an attempt to understand the content; this interpretation is supported by Example 2, which notably features a high number of clarification requests in a very short segment of dialogue.

Interestingly, there were more turn-initial agreement markers in both conditions than in our baseline control condition and fewer turn-initial disagreement markers in our disagreement condition transcripts. This would suggest that the disagreement interventions led to less disagreement and more agreement, as indicated by the markers identified by Walker et al. (2012c). However, it is worth noting that as in the corpus study of the BNC, many of these markers did not appear at all in the dialogues. As such, it seems likely that while they may be characteristic of forum posts they may not feature commonly in real-time dialogue.

The Constructive Disagreement hypothesis also anticipated that the Disagreement condition would lead to more consideration of possible solutions over the course of a dialogue. The results show that although there is an effect of condition on the number of shifts in stance state or solutions considered during the dialogue, the directionality was contrary to our predictions. The insertion of agreement fragments appeared to lead to more shifts in stance. As the intercoder reliability co-efficient was between ($\alpha$= 0.67 - $\alpha$=80), this conclusion remains tentative, but one possible explanation may be that agreement tokens were received as feedback and continuer markers, thus creating a more supportive environment for developing and exchanging possible ideas. The insertion of agreement tokens in the turn initial position could also have transformed a disagreement turn into a politely formulated, agreement prefaced disagreement turn. For example, "but Tom's not very useful" could become "I agree but Tom's not very useful". This would potentially align with the CA and Politeness Theory literature, as well as Chiu (2008), which specifies that while *polite* disagreement may yield more novel contributions, impolite disagreement is always problematic.

Overall, our results most strongly confirm the Politeness hypothesis. Insertions of exposed disagreement had a disruptive effect upon the dialogues, producing confusion and prompting additional clarification requests due to their unexpected and incongruous nature. Conversely, exposed agreement, even though also inserted randomly, did not disrupt the dialogue in the same way and was never explicitly addressed by a participant. The Disagreement condition produced significantly more instances of exposed agreement, which is most easily interpreted in terms of politeness, face and redressive action; with additional exposed disagreement being introduced into the dialogues, it seems that participants respond with cooperation and attempt to redress the potential affronts to face posed by the inserted fragments. As predicted there were more self-edits in the Disagreement condition, suggesting that

participants were having to work harder to respond to the potentially face-threatening insertions. Our results most strongly support the Politeness hypothesis and confirm that exposed disagreement is problematic and disruptive in dialogue.

## 4.5 Implications

The experiment presented in this chapter highlights the problematic effect of exposed disagreement, in particular, how it disrupts dialogues and instigates additional work for participants to coordinate on the discussion process. The introduction of unmitigated disagreement does not serve to advance the deliberative quality of the discussion, rather, it appears to strongly deviate from the expected conversational conduct. In response, to unmitigated disagreement participants issued shorter replies, typed fewer words, which cold be displays or signals of disengagement. Furthermore, they issued more agreement, which suggests that the discursive practice was further compromised by the redressive action which was taken to assuage the face-threat of the disagreements. Dialogues were closed down, shorter answers provided and additional agreement tokens issued more readily in the face of explicit oppositional positioning. Consequently, this suggests that exposed disagreement is not a useful device for fostering constructive engagement; rather, more subtle ways of positioning oppositional content, which are not so obviously face-threatening, although still positioning a speakers contribution as an alternative, should be explored.

# Chapter 5

# Opening Up and Closing Down Discussion: Experimenting with Stance in Conversation

Chapter 4 demonstrated that exposed and unmitigated disagreement can be disruptive and problematic in discussion dialogues. In this chapter the effect of more subtle devices that speakers employ to position their contributions as oppositional are examined.

## 5.1 Introduction

During a debate people have choices about how they present their contributions. Amongst other things they can simply assert their position, they can modify it with a propositional attitude verb such as 'know' or 'think' or they can turn an assertion into a question rephrasing "I think X" as "Do you think X?". These choices of attitude and modality all help to establish what a person's stance is and, in combination with the choices made by their interlocutor, set the tone and direction of a debate. One of the most important hypotheses about the impact of different stance markers on dialogue relates to expressions of epistemic certainty; framing of a stance as 'knowing' or 'unknowing' appears to significantly alter the deliberative quality of a discussion (Heritage, 2012a).

Although the interactional dimensions of stance have been discussed in some detail (Du Bois, 2007; Englebretson, 2007; Kärkkäinen, 2003), this work is based on case studies and corpus analyses. The causal *effects* of adopting different stance markers on the subsequent trajectory of a dialogue has not, as far as we are aware, been directly tested. One key reason for this is the practical difficulty of manipulating stance markers in a live dialogue. In this chapter, we again use the technique

introduced by Healey et al. (2003) that takes advantage of the potential of text-chat for enabling selective manipulation of people's turns, including the addition of stance markers, without their awareness. We use this technique to assess how the epistemic status of a stance, i.e. whether it is framed as either *unknowing* or *knowing*, impacts on the quality of the joint action and deliberation in discussion dialogues.

### 5.1.1 Taking a Stance

Disagreement is generally minimised in conversation (Concannon et al., 2015b; Pomerantz, 1984a), and when executed directly or impolitely tends to be problematic (Chiu, 2008; Concannon et al., 2015a). However, in certain contexts, such as problem solving and discussion tasks, it can be important for advancing the deliberative quality of a dialogue and encouraging novel contributions (Chiu, 2008).

There is thus a delicate balance between mitigating the socially problematic aspects of disagreement while still being able to identify and resolve differences of opinion. This balance can be achieved in many different ways. Resources such as 'well'-prefacing (Pomerantz, 1984a), stance markers such as 'I think' (Kärkkäinen, 2003) and reported speech (Clift, 2006a; Concannon et al., 2015a; Holt and Clift, 2007) all provide less explicit ways of marking what follows as potentially incongruous or in opposition to what went before.

An important underlying function of these different devices is to signal people's claimed rights to speak on a topic. Heritage (2012a) refers to this as 'epistemic status': the relative positioning in which "persons recognize one another to be more or less knowledgeable concerning some domain of knowledge", which can alter from moment to moment, and be "disassembled by persons who deploy epistemic stance to appear more, or less knowledgeable than they really are" (Heritage, 2012a).

There are significant potential social and interactional implications of positioning ourselves or others as either knowing or unknowing (Levinson, 2012). The management of rights and responsibilities is closely connected to participants' concerns with 'face' (Heritage and Raymond, 2005) or the "positive social value a person effectively claims for himself" through interaction (Goffman, 1967a). For example, in issuing requests for information or questions, the requester assumes an unknowing epistemic status, and positions the recipient in a knowing one (Heritage, 2012a). Such requests also create an obligation for the recipient to respond to the requester (Levinson, 2012). Levinson (2012) observes that people prefer polar questions to other forms that require more knowledge-rich responses and often disguise them as assertions, thus demonstrating an unwillingness to locate oneself in an unknowing position, nor to impose too greatly upon an interlocutor by demanding a response.

### 5.1.2 *Knowing* vs *unknowing* stances

Following Heritage (2012b) (addressed in detail in section 2.5.3) our general hypothesis is that framing a proposition as *unknowing* invites elaboration, sequence expansion and further discussion of the topic at hand (Heritage, 2012b). Conversely, a more *knowing* epistemic stance, creates a pressure for confirmation and sequence closing. As such, we predict that inserting 'knowing' and 'unknowing' stance markers will have different impacts on the course of a conversation even where, counterfactually, nothing about the content of the modified assertions is changed.

Heritage (2012a) defines 'epistemic status' as the relative positioning in which "persons recognize one another to be more or less knowledgeable concerning some domain of knowledge". Knowing all (K+) is typically conveyed through declaratives, while interrogative grammatical format is the most explicit way that a speaker can embody an 'unknowing' (K-) epistemic status. For example, the question 'what time is your appointment' positions the speaker in request of information, where as 'your appointment is at 3pm' positions the speaker in a K+ position. However, as highlighted by Drew (2012), how much speakers know relative to one another is not only encoded in the grammatical format, but also in incongruities between epistemic status and grammatical format, for example in posing a question to which you already know the answer (e.g. 'Aren't you going to be late?'). Speakers' relative positioning can alter from moment to moment, and be "disassembled by persons who deploy epistemic stance to appear more, or less knowledgeable than they really are" (Heritage, 2012a).

In issuing a question the requester assumes an unknowing epistemic status and positions the recipient in a knowing one (Heritage, 2012a), creating an obligation for the recipient to respond (Levinson, 2012). Levinson (2012) observes that people prefer polar questions to other forms that require more knowledge-rich responses and often disguise them as assertions, thus demonstrating an unwillingness to locate oneself in an unknowing position, nor to impose too greatly upon an interlocutor by demanding a response. However, in a discussion context, in which individual contributions on the topic under discussion are warranted and expected, the ways in which requests are made could be influential to the deliberative quality of the discussion.

Furthermore, between the most explicit formats of K+ and K- constructions (i.e. declaratives and interrogatives), there are a range of other ways that speakers can encode epistemic stance, such as modals, hedges and epistemic adverbs, which can convey levels of speaker certainty, e.g. 'It was definitely red', and commitment 'I absolutely think...' and evidential markings which convey the source of a knowledge claim (i.e. direct evidentials based on sensorial/ visual evidence and

indirect evidentials, such as inference and hearsay). Particularly within a discussion context the management of imbalances in epistemic status is particularly pertinent as participants' contributions must necessarily negotiate alternative stance positions.

### 5.1.3 Collaborative and individual stance marking

In group work and decision-making which involves negotiation, the distinction between individual and collective stance taking becomes increasingly important. While stance-taking has been framed as inherently interactional (Du Bois, 2007; Englebretson, 2007), and the the negotiation of stance can indeed be highly interactive, the ways in which stance is marked can be individual or collective, i.e. a person can present a stance as their own, or jointly shared with other conversational partners. One obvious way in which this effect is achieved is through the use of personal pronouns, with 'I' typically used to represent a personally claimed stance, and 'we' used to propose a joint stance. The use of 'you' can either be used to demarcate the perceived opinion of an interactional interlocutor or serve as an interactional tool. For example, Landgrebe (2012) addresses the function of the epistemic markers 'I think' and 'you know', identifying the function of the latter interactional resource as "to orient to a common understanding and invite for involvement and a shared epistemic stance" (Landgrebe, 2012). Consequently, an additional hypothesis is that framing propositions as 'unknowing' will also lead to greater collaboration and the production of joint stances.

## 5.2 Marking Stance in the Balloon Task

The task chosen for the experiment reported below is the Balloon Task. Participants are presented with a fictional scenario in which an hot air balloon is losing altitude and about to crash. The only way for any of three passengers to survive is for one of them to jump to a certain death. The three passengers are: Dr. Nick Riviera, a cancer scientist, Mrs. Susie Derkins, a pregnant primary school teacher, and Mr. Tom Derkins, the balloon pilot and Susie's husband. To ensure a relatively natural manipulation of epistemic stance was selected for this task an initial analysis was conducted using control condition transcripts from previous balloon task transcripts generated in the experiment presented in chapter 4. Twelve transcripts were analysed for markers that conveyed 'knowing' or 'unknowing' states in relation to stance marking. 'I think' was frequently used as a resource to mark a stance position. 'I think' has been attributed a dual function, and can can also act as a hedge (Holmes, 1990), however in the discussion context it was used most frequently to convey a knowing stance, particularly when at the beginning of a turn.

(1) a. *I think* Tom should definitely stay in the balloon

b. *I think* Nick should definitely be the one to go

c. *I think* because there's an element of risk with whether Nick will actually end up coming up with a cure for cancer ... There's no point taking two risks by then letting go of Tom

d. *i think* we have a couple mins left

e. A: so tom has to jump?
B: *i think* so

In 1a, 1b and 1c the marker 'I think' serves to accentuate the propositional content and emphasise the speaker's commitment to their proposition and focuses on a substantive aspect, namely, who should be sacrificed. In 1d and 1e, however, the marker performs the opposite effect and suggests a lack of speaker commitment and acts as a hedging marker. There were 44 instances of 'I think' in the transcripts, 34 instances (77.27%) served to emphasise the propositional content it was associated with, eight instances (18.18%) acted in a 'hedging' or *unknowing* capacity, and the two remaining instances made manifest the cognate processes (e.g."whenever i think that nick should go, i think 'Are susie and tom really that important?'").

Of the 44 instances, 25 were turn-initial (56.82%), 19 of which served to emphasise the speaker's ownership and commitment to the content that followed. Four instances of turn-initial 'I think' (20%) were constructed in such a way that 'I think' functioned as a hedging marker and two instances were not possible to classify due to insufficient context (i.e. 'I think overall'). Closer inspection of the use cases showed that all instances of turn-initial 'I think', in which the proceeding content featured a character from the scenario, conveyed a 'knowing' stance. As such, using 'I think' as a turn-initial insertion for turns which contain a mention of one of the scenario's characters, should increase the likelihood of a consistent effect of framing the utterance as 'knowing', rather than performing a hedging effect.

When used non-turn-initial we see more variation in use - so whether as a hedge or booster. For example in the following example:

A: im sure plenty of doctors think that they're on the brink
B: Yeah
A: he is very smart, so he probably isn't over exaggerating
A: he might be but i think its unlikely that he is exaggerating

While looking for markers which served to downgrade the epistemic strength of assertion, 'do you think' was one such 'unknowing' device that was used in the transcripts.

> A: do you think the married couple would gang up on the doctor and throw him out
> B: maybe. he is their friend though

'Do you think' makes a minimised contrast pair with 'I think' and can be inserted at a turn-initial position without changing the content of the turn. Consequently, 'I think' and 'do you think' were selected as our turn-initial inserts to frame the proceeding content as more or less 'knowing'.

## 5.3 Hypothesis

The analysis of previous dialogues enables us to operationalise our general hypothesis about the level of knowingness with which opinions are presented. In particular, inserting 'I think' should, all things being equal, position assertions as more 'knowing', i.e. frame them as a stance with increased speaker commitment. If this is correct it should tend to cause a closing down of the discussion of the topic at hand. As such we anticipate there will be less deliberation about the possible solutions to the task. Conversely, we expect that presenting opinions as 'unknowing' should improve the deliberative quality of the dialogue, with more ideas and positions exchanged and elaborated on. We anticipate that presenting contributions as unknowing will create an environment in which participants are more likely to make manifest their uncertainty, as it will be acceptable to appear unknowing, *and* certainty, as solutions will be discussed more, and potentially co-constructed, so that once a stance is established it can be committed to with greater conviction.

## 5.4 Predictions

1. Fewer possible solutions will be considered when contributions are framed as knowing and responses will be less considered; this should affect turn formulation, with shorter typing times and less editing of turns. Framing contributions as knowing will close down the dialogue, as indicated by shorter and fewer turns.

2. Framing contributions as unknowing will open up dialogues, leading to longer turns and more possible solutions considered.

3. More possible solutions will be considered and more care will be taken in the construction of turns, as evidenced by slower typing times and more edits when contributions are introduced with an unknowing preface (‘do you think X’).

4. Framing contributions as unknowing will lead to higher frequencies of certainty and uncertainty markers.

## 5.5 Method

Pairs of participants were seated at separate computers on either side of an open plan office and given an instruction sheet to read detailing the balloon task. The task presented is identical to the one used in Chapter 4. Participants are told to take as much time as they need to read the summary of the situation and then discuss with their partners via a chat tool set up on the computer at which they are seated, and attempt to come to a conclusion over who should jump from the balloon. The advantages of this task are that it is effective at generating debates between subjects and there is good scope for deliberation.

### 5.5.1 Participants

Sixty participants were recruited, 41 females and 19 males, with the majority being undergraduate and postgraduate students at the University of London. Participants were invited to attend with someone who they already knew. They were recruited in pairs to ensure that inter-pair participants were acquainted and increase the likelihood that both participants would attend. For a couple of experiments participants partnered up with someone else from their course, but with whom they were not previously acquainted, but these exceptions are marked in the data. All subjects were native speakers of English. Each participant received 7.50 GBP remuneration for participating in the experiment, or if they were a Psychology student at Queen Mary University of London, they could receive course credits in lieu of payment.

### 5.5.2 Design

The experiment is conducted in pairs, with 10 dyads per condition. Pairs of participants were presented with a discussion task and instructed to discuss for 30 minutes and attempt to come to an agreement. The three experiment conditions were:

1. Control Condition: No interventions

2. Knowing Condition: Turn-Initial ‘I think’ Insertion

3. Unknowing Condition: Turn-Initial ‘Do you think’ Insertion

Each pair of participants was assigned to one of these conditions at random. In the Control condition there were no interventions performed by the server; participants received the dialogue turns exactly as they were typed. However, in the Knowing and Unknowing conditions manipulations were carried out. Interventions were issued every four turns, if and only if the turn included a reference to one of the characters in the scenario (e.g. Doctor, Susie, etc.). Interventions are not visible to the individual whose turn has been manipulated, only the recipient, so that there is no awareness that turns are being intercepted before being relayed.

To increase the consistently of our ‘I think’ intervention functioning in a more knowing capacity we created a rules based system that ensured that a consistent effect was produced by the insertion- as far as possible. To Control for unnecessary variance we introduced some protocols to ensure that an intervention would only be triggered if the utterance featured a subject/object. Statements about the characters within the scenario, and which did not functioned as a question, were targeted.

#### Frequency of interventions

A pilot study was conducted to assess how frequently an intervention would need to be inserted. It was important that we could achieve enough interventions, but without creating clearly nonsensical statements. Making insertions every four turns led to enough interventions over the course of a dialogue so that however, in some cases this led to poorly formulated turns being produced. It was decided that the intervention would take place every four turns, if and only if the sentence contained reference to a character (i.e. variations of dr/Nick/scientist, balloon pilot/Tom or Susie/Sue, he/she/they etc.). This was achieved using a keyword search on turns before they were relayed to the participant, which occurred in real time to ensure there was no perceivable delay. The conversation Controller was written in such a way to ensure that an insertion was only produced if the utterance that followed was *not* a question, a polar answer (i.e. sentences starting with a turn-initial yes or no), comprised of a single word, nor started with a continuer, i.e. ‘and’.

### 5.5.3 Procedure

Upon arrival, I personally met all participants and provided an overview of the study, allowing for any questions to be asked. After completing consent forms and discussing any queries that they had I took each participant to a separate computer. The experiment took place in an open plan office; participants were sat at computers in the same room but at other sides of the office. As it was an open plan office

there were often other colleagues walking in and out or working quietly in the room. Participants were made aware that this may happen before the experiment started, so that they were not distracted. Participants were informed at the start of the experiment that they were free to leave at any time if they felt uncomfortable, however, there were no instances of this happening. Before starting the task participants are given the information sheet (see chapter A), which outlines the scenario. Once they had finished reading the instruction sheet they were free to start typing. After 30 minutes I asked them to stop typing and they were provided with a brief questionnaire (see chapter A) to complete, before a full debrief was conducted and remuneration handed over.

### 5.5.4 Analysis

The DiET chat tool records all interventions and key presses, including edits made before participants press 'send', made by each participant. Counting frequencies of certainty adverbials (e.g. 'surely', 'clearly', etc.), uncertainty markers ('possibly', 'potentially', etc.) and personal pronouns ('we/us' as indexing collective stance, 'I/me/my' for individual stance, and 'you/yours' for other stance) were also collected.

Uncertainty Markers includes uncertainty adverbials as well as modals ('may', 'might' and 'could') and hedges ('quite', 'sort of', etc.), but certainty and uncertainty adverbials are also presented individually for comparison[1] . Obvious typographical errors were corrected to increase the accuracy of the frequency counts (e.g. possibiliyt –> possibility). The inserted fragments were also removed from the transcripts before frequency counts were conducted, to ensure that the figures reflected only what the participants actively contributed.

### 5.5.5 Coding for shifts in stance position

As in the previous experiment it was necessary to assess how the experimental conditions affected the levels of deliberation in the dialogues.

In order to assess the deliberative quality of each dialogue, two factors were considered: the number of possible solutions considered and the total number of **stance positions** inhabited during the dialogue. A shift from one stance position to another is referred to as a **stance shift**. The number of **alternative solutions**

[1]The full list of words and phrases used for each category are as follows:- Certainty adverbials: absolutely, actually, certainly, clearly, plainly, definitely, evidently, indeed, obviously, really, surely, undoubtedly, unquestionably, for certain, for sure,of course; Uncertainty adverbials: admittedly, allegedly, apparently, arguably, conceivably, inexplicably, likely, maybe, perhaps, possibly, potentially, presumably, probably, reportedly, seemingly, supposedly; Uncertainty modals: may, might, can could Hoye (1997) and Hedges (including approximators): quite, sort of, kind of , might, a bit, a little bit, just, at least, approximately, about, around, something like, almost, pretty, sometimes.

**considered** refers to the total number of possible solutions entertained (out of a total four possible solutions). This serves as a measure for the quality of the deliberation, and demonstrates how much of the state space of the problem solving domain they explored. The dialogue transcripts are hand coded to record this information and indicate points in the dialogue when the two participants share the same or opposing stances.

The transcripts were hand coded for shifts from one stance position to another regarding who to throw off of the balloon, i.e when a participant conveyed through the text a consideration of a solution regarding who to sacrifice. The stance positions were hand labelled for each conversation so that it is possible to detect the decision patterns for each participant, how it relates to their conversational partner, and to gain a numeric indicator of how many times the participant shifted their stance position. The transcripts were hand coded for the solutions being considered, with a possible solution being the character to be sacrificed (for example, Undecided, Kill Tom, Kill Susie, Kill Nick) and the number of shifts from one solution to another during the conversation (e.g. Kill Susie –> Undecided –> Kill Tom = three shifts in stance position).

The coding scheme used for labelling the stance positions was a simplified version of the protocol used in the previous experiment. The decision to alter the coding method was taken to reduce the level of annotator interpretation necessary for coding, simplifying the number of factors considered to make the coding process less ambiguous and robust. Due to the nature of the task, participants often would take a structured approach and consider each of the inhabitants of the balloon in turn; consequently this led to additional shifts in stance being recorded that did not necessarily reflect a committed viewpoint, as deliberations from someone being definitely saved, to possibly not could be marked as a shift, but required additional annotator interpretation. In this experiment the coding only marked a shift in the decision of who to sacrifice, rather than who to save *and* who to sacrifice. Consequently, the number of possible stance states was reduced from seven to four. The three removed stance positions were:

1. Save Susie but undecided on who should die;
2. Save Nick but undecided on who should die
3. Save Tom but but undecided on who should die

The four remaining potential stance positions were:

1. Undecided
2. Sacrifice Susie (and therefore save the other two)

3. Sacrifice Nick (and therefore save the other two)

4. Sacrifice Tom (and therefore save the other two)

The point in the dialogue where each of these stance position shifts occur was also recorded for each participant. The annotations are based purely on the content of the dialogue and when it is made manifest in an utterance that a participant has shifted opinion, even though the participant may have formed an opinion before they share it with their partner via text. Furthermore, turns were counted in which participant A and B had matching or opposing stance states. Time spent in an undecided state, even if both participant A and B both were undecided is not counted as a matching stance as it is unclear what their current stance is. A matching stance would only be A: Tom B: Tom; A: Nick B: Nick or A: Susie B: Susie.

#### Intercoder reliability

The annotation for stance position and shifts in stance position was done by a single annotator, blind, and all labelling indicating which condition a file belonged to was removed. A participant's stance position was carried over to the next turn, unless their current turn provided new information that contradicted the previous stance position, in which case the stance position was updated.

A second annotator coded a sub-sample (approximately 10% of the total data) of the experiment data, a random selection of three experiment transcripts across all conditions, comprising 563 turns of dialogue. Inter-annotator agreement was calculated using Krippendorff's alpha (for nominal data) and the intercoders met intercoder reliability ($\alpha$= 0.85).

See Section A.4 for details of the guidance provided to the second annotator.

## 5.6 Results

In the following sections the results of the experimental analysis is presented.

### 5.6.1 Message construction

#### Word counts

Table 5.1 details the number of total words, total turns and average words per turn per participant for each condition.

| Condition | | Total Words | Total Turns | Mean Words Per Turn |
|---|---|---|---|---|
| Control | Mean | 618.30 | 81.05 | 8.71 |
| | Std. Dev. | 182.12 | 35.91 | 4.43 |
| Unknowing | Mean | 641.95 | 74.35 | 9.45 |
| | Std. Dev. | 160.28 | 34.57 | 3.62 |
| Knowing | Mean | 649.05 | 94.65 | 6.65 |
| | Std. Dev. | 196.08 | 37.50 | 1.87 |
| Total | Mean | 636.43 | 83.35 | 8.27 |
| | Std. Dev. | 177.51 | 36.41 | 3.62 |

Table 5.1 Word counts per condition, per participant

Figure 5.1 shows a boxplot of the mean word counts per condition.

Fig. 5.1 Boxplot of Average Total Words per Dyad per Condition

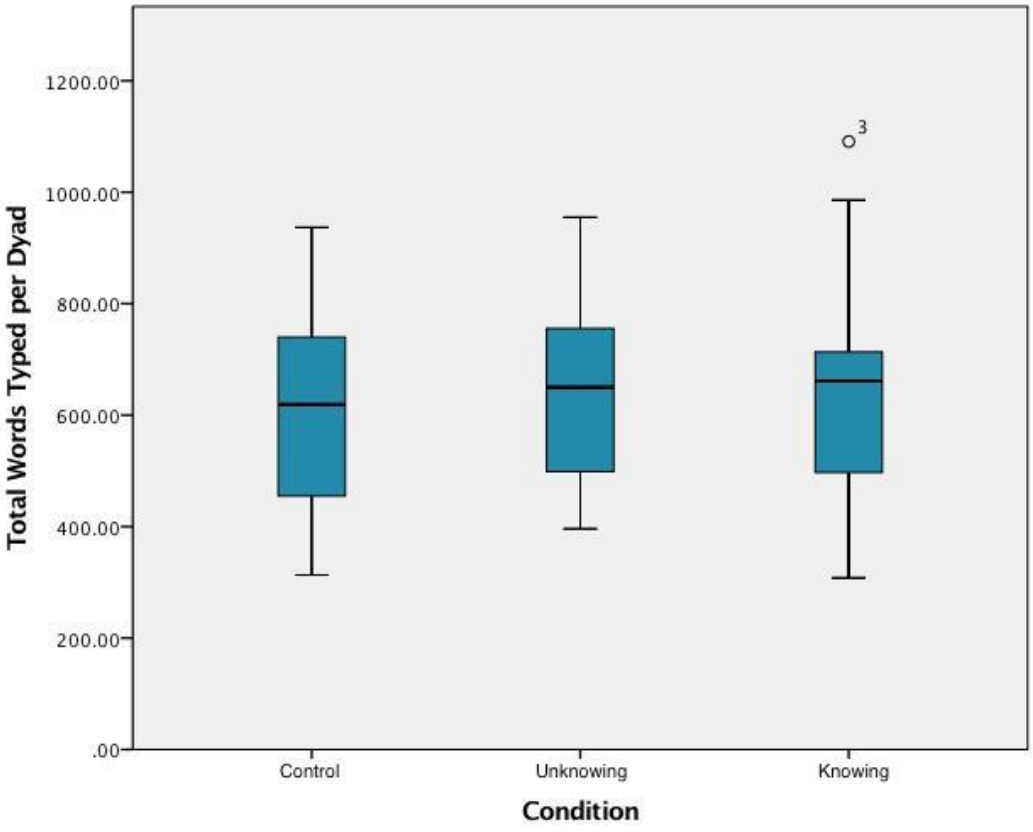



Figure 5.2 is a boxplot detailing the mean number of turns in a dialogue for each condition.

Fig. 5.2 Boxplot of Average Total Number of Turns per Dyad per Condition

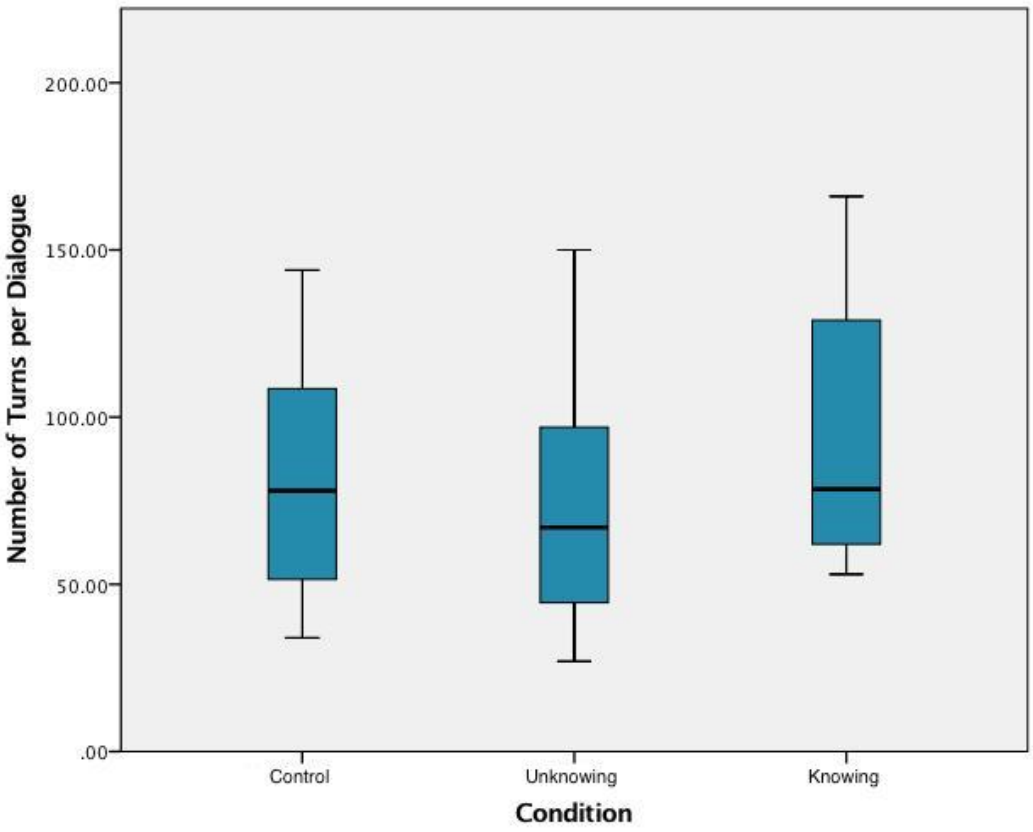
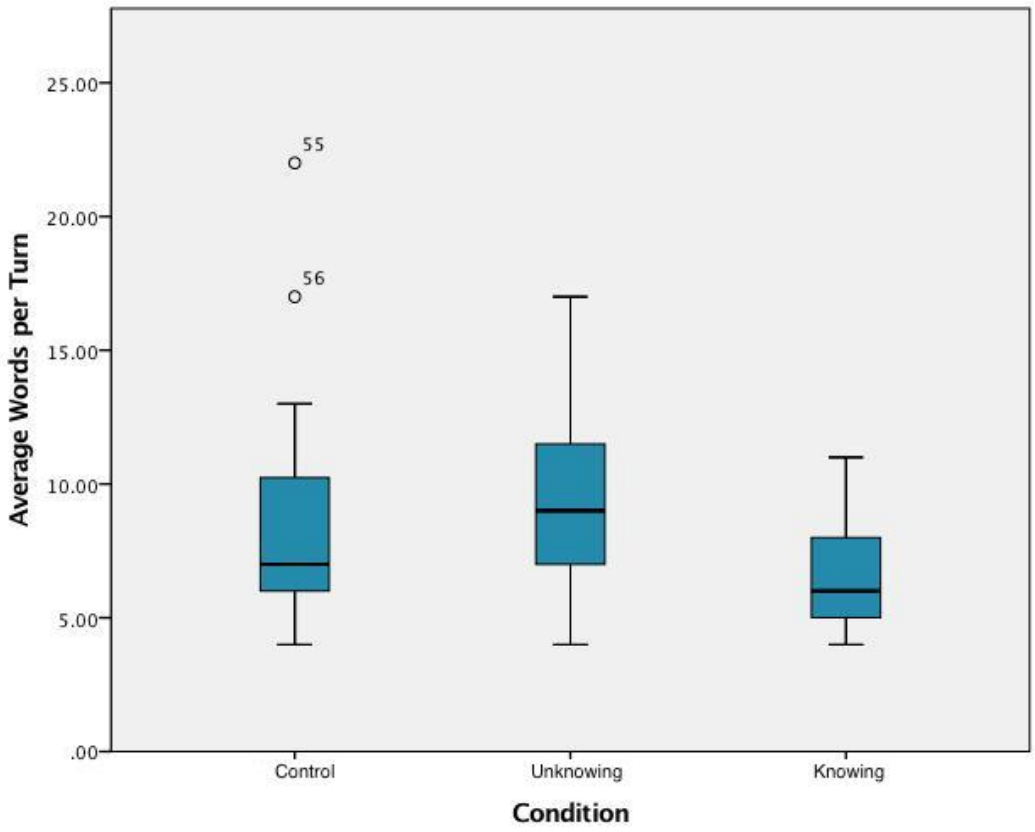



Figure 5.3 is a boxplot detailing the mean number of words in a given turn for each condition in a dialogue for each condition.

Fig. 5.3 Boxplot of Total Average Words per Turn per Condition




A nonparametric Kruskal Wallis independent samples test shows that there is a significant omnibus effect of condition on the number of words typed per turn ($H_{(2)}$ = 7.475 , p= 0.02). A post hoc pairwise comparison using Dunn's test shows that there is a significant difference in the number of words produced in a turn between the Knowing and Unknowing condition, with Knowing dialogues containing fewer words per turn and Unknowing dialogues containing more words per turn (p= 0.02). There is no significant difference between the number of words typed per turn in the Control and Unknowing conditions (p= 0.74), nor Knowing and Control conditions (p= 0.35). A nonparametric Kruskal Wallis independent samples test shows that the total number of words typed was not significantly affected by the condition ($H_{(2)}$ = 0.283, p= 0.87) and there is no significant effect of condition on the number of turns per dialogue ($H_{(2)}$ = 3.556 , p= 0.17).

### Typing time

Table 5.2 provides details of the typing time in milliseconds and the speed of typing for each condition.

| Condition | | Avg. Type Time | Avg. Speed |
|---|---|---|---|
| Control | Mean | 16005.50 | 3.09 |
| | Std. Deviation | 8118.06 | 0.88 |
| Unknowing | Mean | 17279.88 | 3.72 |
| | Std. Deviation | 10793.22 | 1.17 |
| Knowing | Mean | 12988.29 | 3.26 |
| | Std. Deviation | 5300.58 | 0.81 |
| Total | Mean | 15424.56 | 3.36 |
| | Std. Deviation | 8 430.84 | 0.99 |

Table 5.2 Typing Time and Speed per condition, per participant

Figure 5.4 shows the boxplot for the mean duration spent constructing a message across conditions.

Fig. 5.4 Boxplot of Average Typing Time for Message Construction per Condition

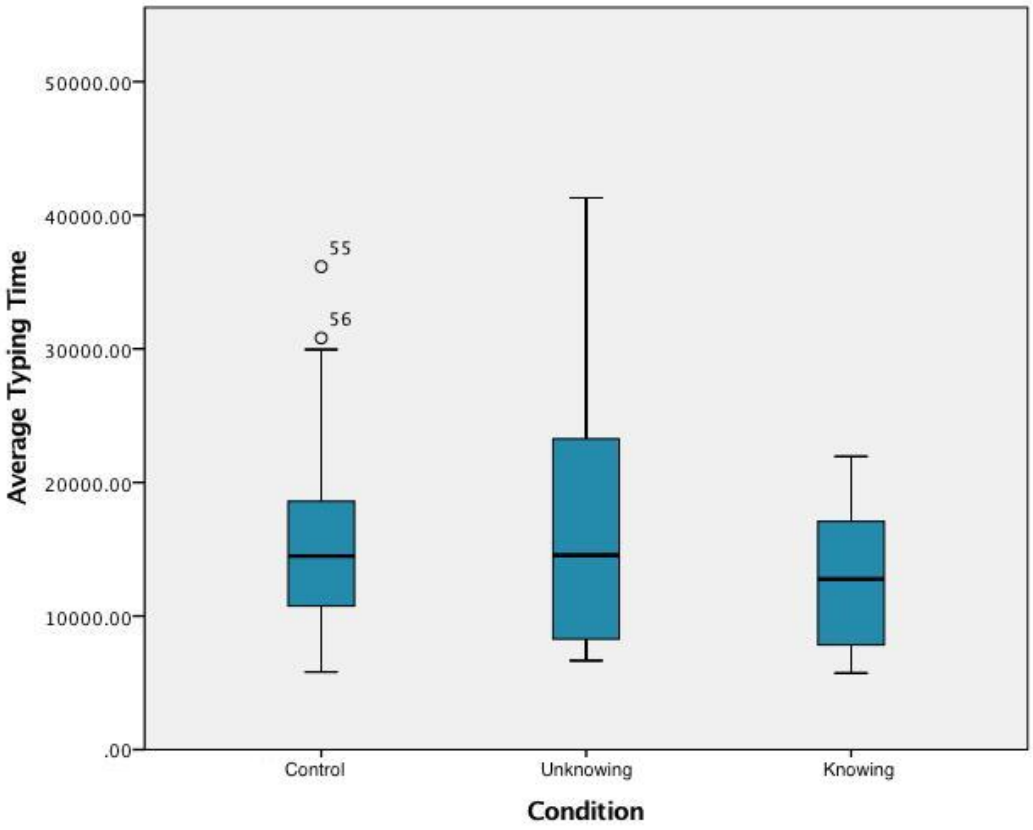
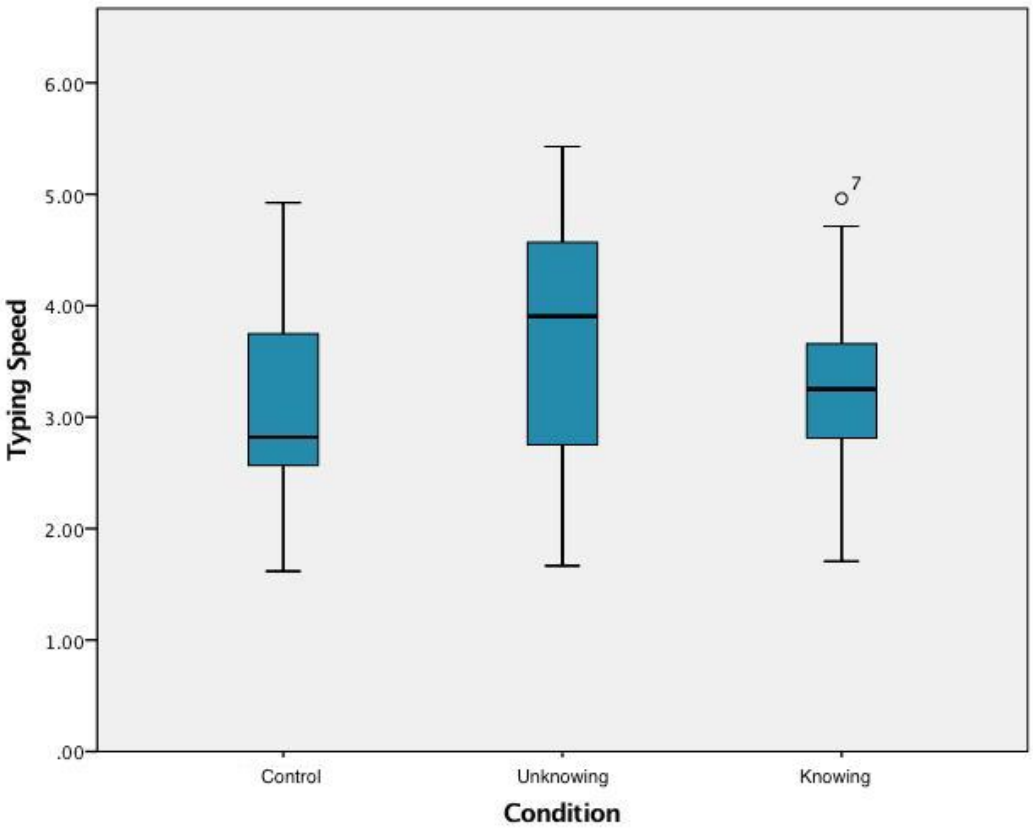



Figure 5.5 displays the boxplot for the mean typing speed of participants across conditions.

Fig. 5.5 Boxplot of Average Typing Speed per Condition




Typing time averaged by participant was analysed using a Generalised Linear Mixed Models analysis (GLMM) with a Gamma distribution because the timing data was positively skewed. Participants was included as a random factor and condition as a fixed factor. This shows a clear main effect of condition ($F(2,59)=13.18$, $p<0.00$). The estimated marginal means are: Control: 12,139, Unknowing: 13,404 and Knowing: 8,813. Pairwise Contrasts show that the Knowing condition has shorter typing times than Control ($t = -3.606$, $p<0.00$) and shorter than the Unknowing condition ($t = -4.87$, $p<0.00$) but Unknowing and Control are not reliably different ($t = 1.16$, $p=0.25$).

### Self edits

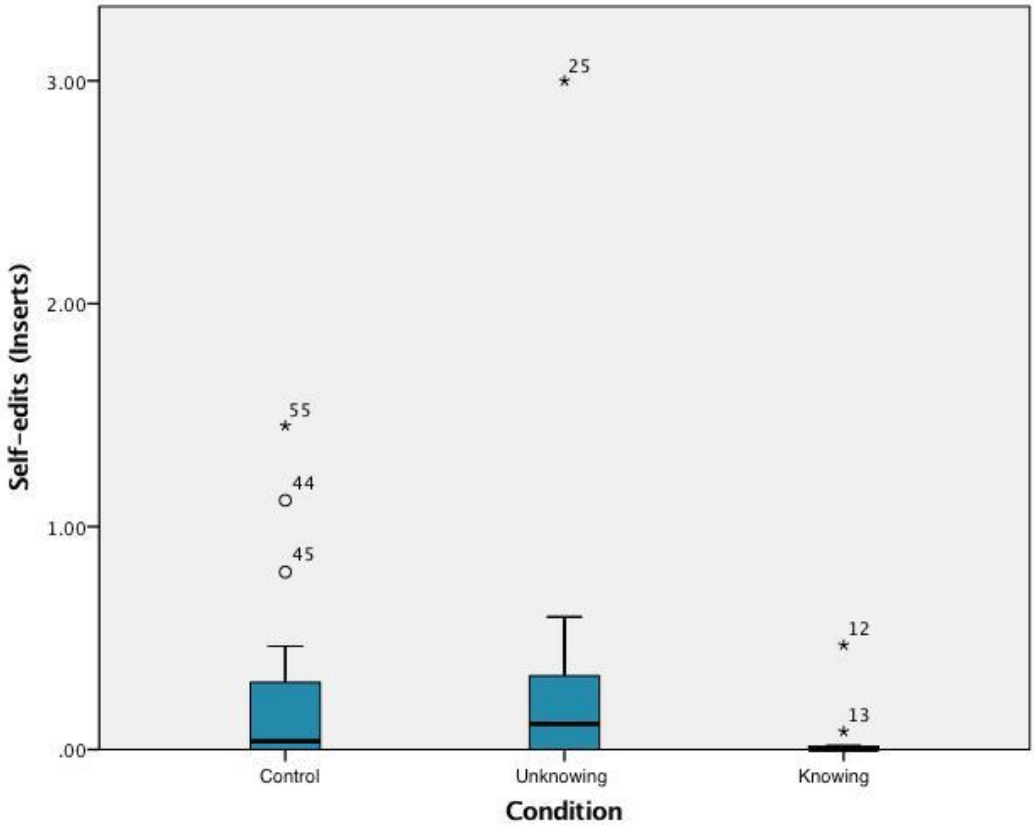
Table 5.3 provides details of the various self-edits participants made during turn construction, such as deletion and insertion of characters before pressing send to relay the message to their partner.

| Condition | | Self-edits (Inserts) | Self-edits (Deletion) |
|---|---|---|---|
| Control | Mean | 0.25 | 53.32 |
| | Std. Deviation | 3.79 | 125.10 |
| Unknowing | Mean | 0.31 | 73.38 |
| | Std. Deviation | 0.68 | 169.25 |
| Knowing | Mean | 0.03 | 20.41 |
| | Std. Deviation | 0.10 | 49.01 |
| Total | Mean | 0.20 | 49.01 |
| | Std. Deviation | 0.47 | 118.79 |

Table 5.3 Self edits per condition, per participant

Figure 5.6 displays the relative distributions of self-edits (inserts) during message construction as a boxplot.

Fig. 5.6 Boxplot of Average Self-edits (inserts) per condition

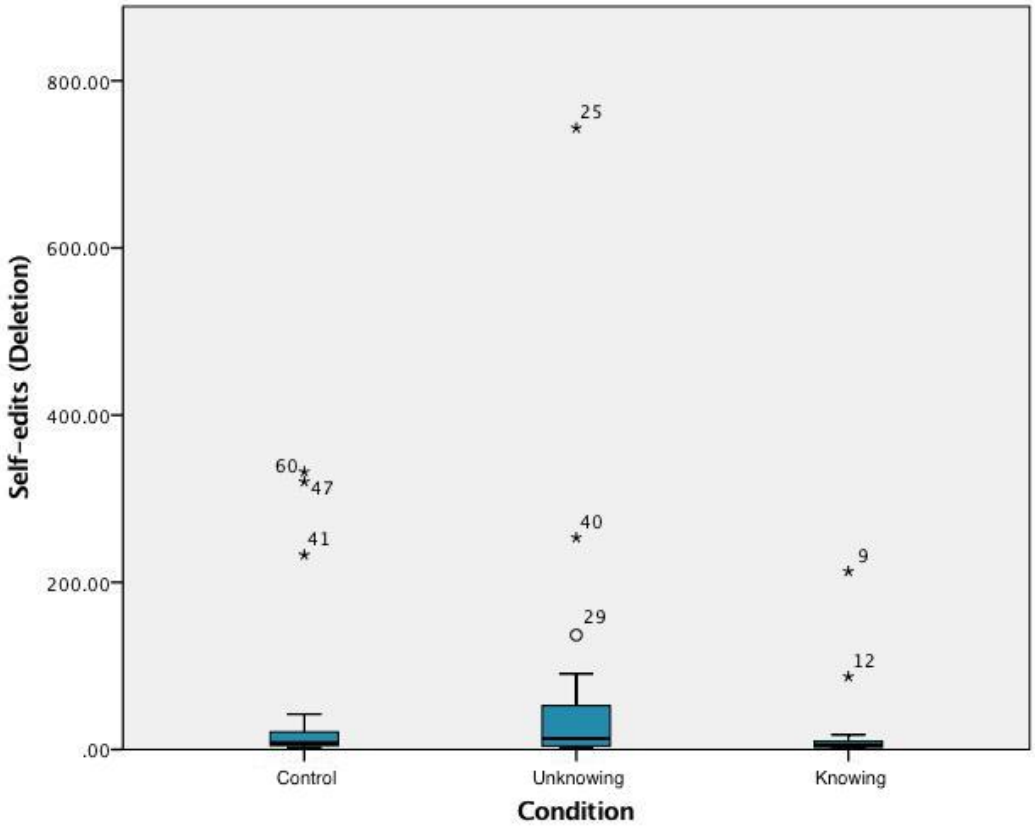



The mean number of self-edit insertions per turn is substantially lower in the Knowing condition than the Control and Unknowing conditions. A Kruskal Wallis test shows that there is a significant omnibus effect of condition on the number of Self-edits (Inserts) per participant ($H_{(2)}$ = 7.761, p=0.02). A post hoc pairwise

comparison using the Dunn's method shows that there are significantly fewer Self-edits (Inserts) in the Knowing condition than the Unknowing condition (p=0.04), but no significant difference between Knowing and Control (p=0.06), nor Unknowing and Control condition (p=1.0). The mean number of self-edit deletions per turn is higher in the Unknowing condition than the Control and Knowing conditions (see fig. 5.7). However, a non-parametric Kruskal Wallis test shows that there is no significant effect of condition on the number of Self-edits (Deletions) ($H_{(2)}$ = 4.560, p=0.10).

Figure 5.7 displays the relative distributions of self-edits (deletes) during message construction as a boxplot.

Fig. 5.7 Boxplot of Average Self-edits (deletion) per condition




## 5.6.2 Epistemic Strength

Table 5.4 provides mean frequencies of epistemic markers, adverbials of certainty, adverbials of uncertainty and combined uncertainty markers (adverbials, hedges, modals) per 100 words.

| Condition | | Certainty Adverbials | Uncertainty Adverbials | Uncertainty Markers |
|---|---|---|---|---|
| Control | Mean | 0.28 | 0.54 | 4.69 |
| | Std. Dev. | 0.25 | 0.32 | 1.12 |
| Knowing | Mean | 0.33 | 0.55 | 4.60 |
| | Std. Dev. | 0.34 | 0.21 | 1.19 |
| Unknowing | Mean | 0.67 | 0.65 | 4.69 |
| | Std. Dev. | 0.35 | 0.39 | 0.88 |
| Total | Mean | 0.43 | 0.58 | 4.66 |
| | Std. Dev. | 0.35 | 0.31 | 1.04 |

Table 5.4 Epistemic markers

A non-parametric Kruskal Wallis test shows that there is an omnibus effect of condition on the frequency of certainty adverbials ($H_{(2)}$ = 7.501 p=0.02). A post-hoc pairwise comparison Dunn's test shows that there are significantly more certainty adverbials in the Unknowing condition compared to the Control condition (p=0.04), but no significant difference in frequencies between the Control and Knowing (p=1.00), nor Knowing and Unknowing conditions (p=0.08). A non-parametric Kruskal Wallis test shows that there is no omnibus effect of condition on the mean frequencies of uncertainty adverbials ($H_{(2)}$ = 0.742 p=0.69) or combined uncertainty markers ($H_{(2)}$ = 0.148 p=0.93).

Figure 5.8 displays a boxplot showing the frequencies of certainty adverbs across conditions.

Fig. 5.8 Boxplot of Certainty Adverbials

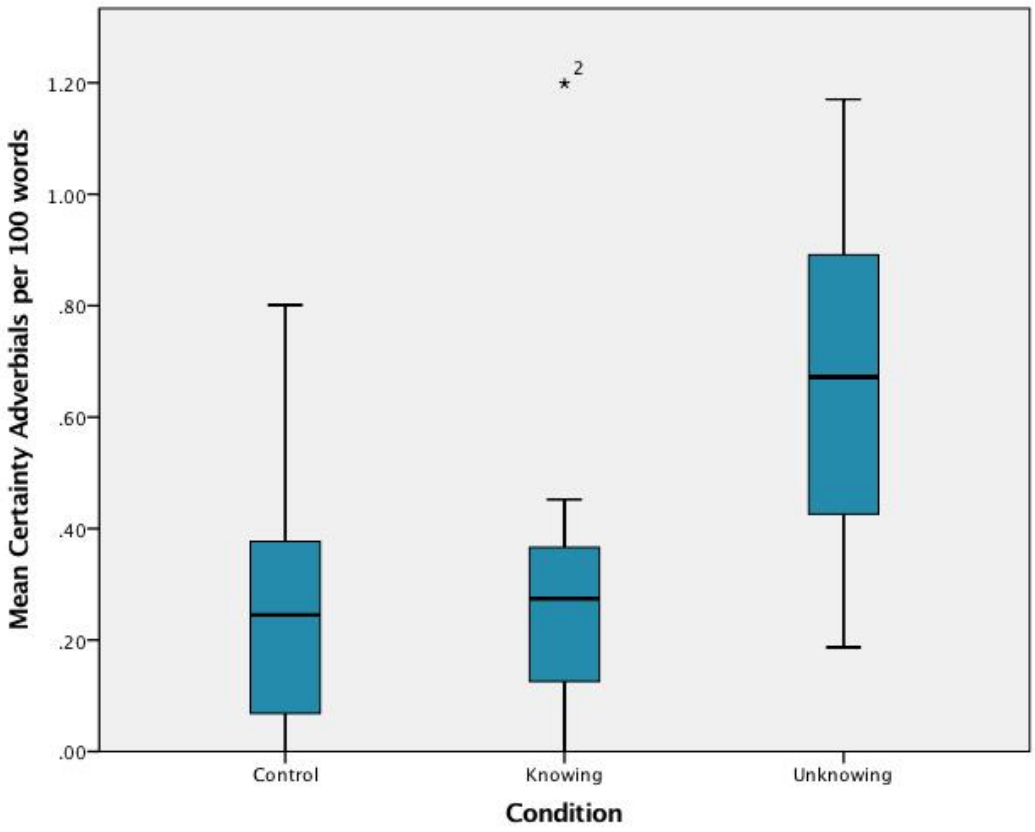



Figure 5.9 displays a boxplot showing the frequencies of uncertainty adverbials across conditions.

Fig. 5.9 Boxplot of Uncertainty Adverbials

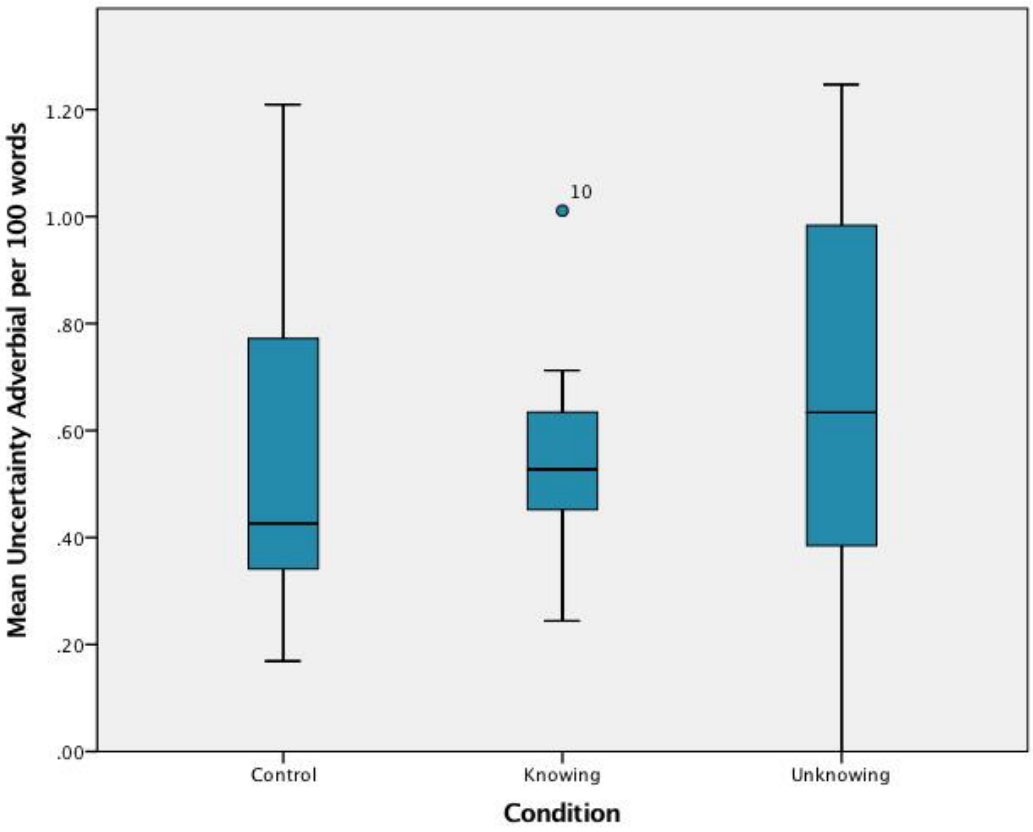
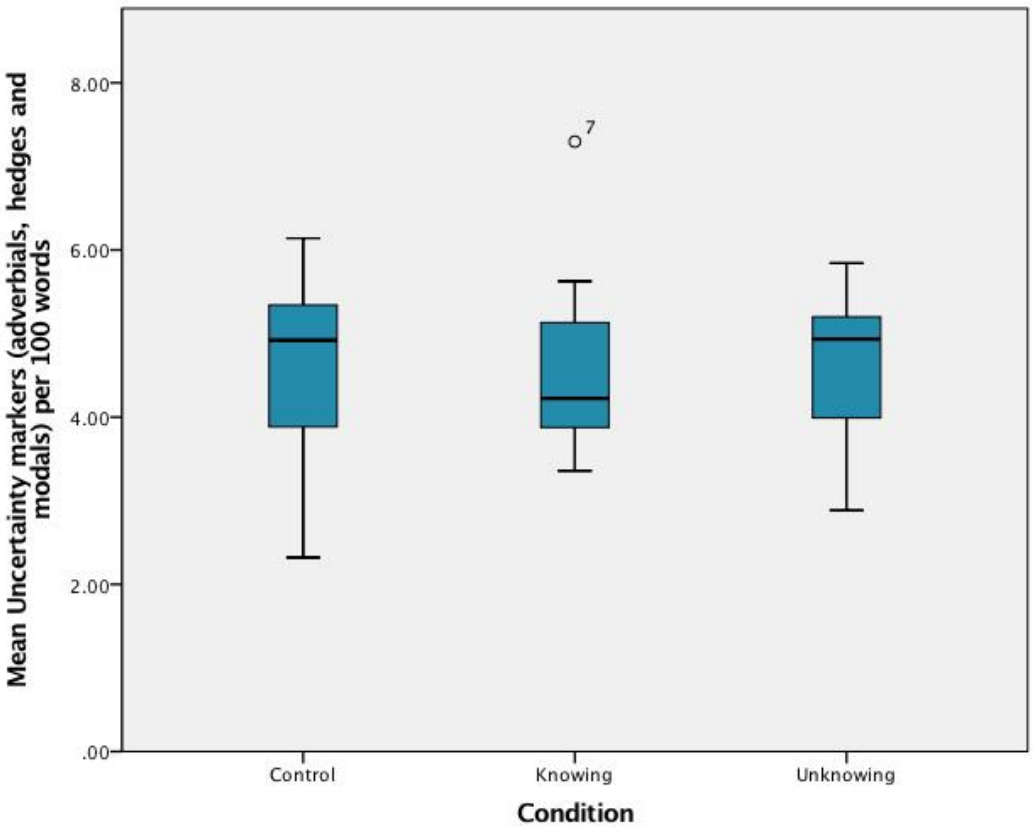



Figure 5.10 displays a boxplot showing the frequencies of combined uncertainty markers across conditions.

Fig. 5.10 Boxplot of Combined Uncertainty Markers




### 5.6.3 Markers of Contrast and Negation

Table 5.5 provides details of the mean frequencies of contrast and Negation Markers normalised per hundred words, for each condition.

| Condition | | Contrast Markers | Negation |
|---|---|---|---|
| Control | Mean | 1.54 | 1.04 |
| | Std. Deviation | 0.69 | 0.32 |
| Unknowing | Mean | 1.69 | 1.52 |
| | Std. Deviation | 0.63 | 0.41 |
| Knowing | Mean | 1.51 | 1.38 |
| | Std. Deviation | 0.55 | 0.80 |
| Total | Mean | 1.64 | 1.31 |
| | Std. Deviation | 0.64 | 0.57 |

Table 5.5 Contrast and Negation

Figure 5.11 displays a boxplot showing the frequencies of negations across conditions.

Fig. 5.11 Boxplot of Negation Frequencies

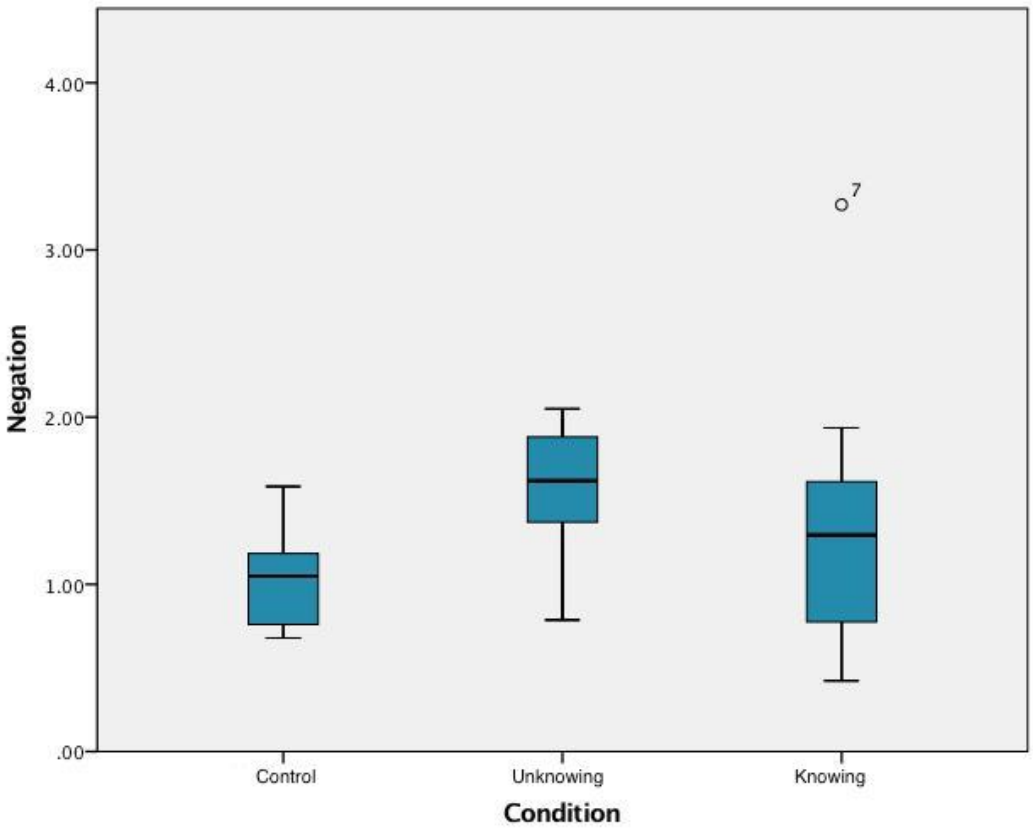


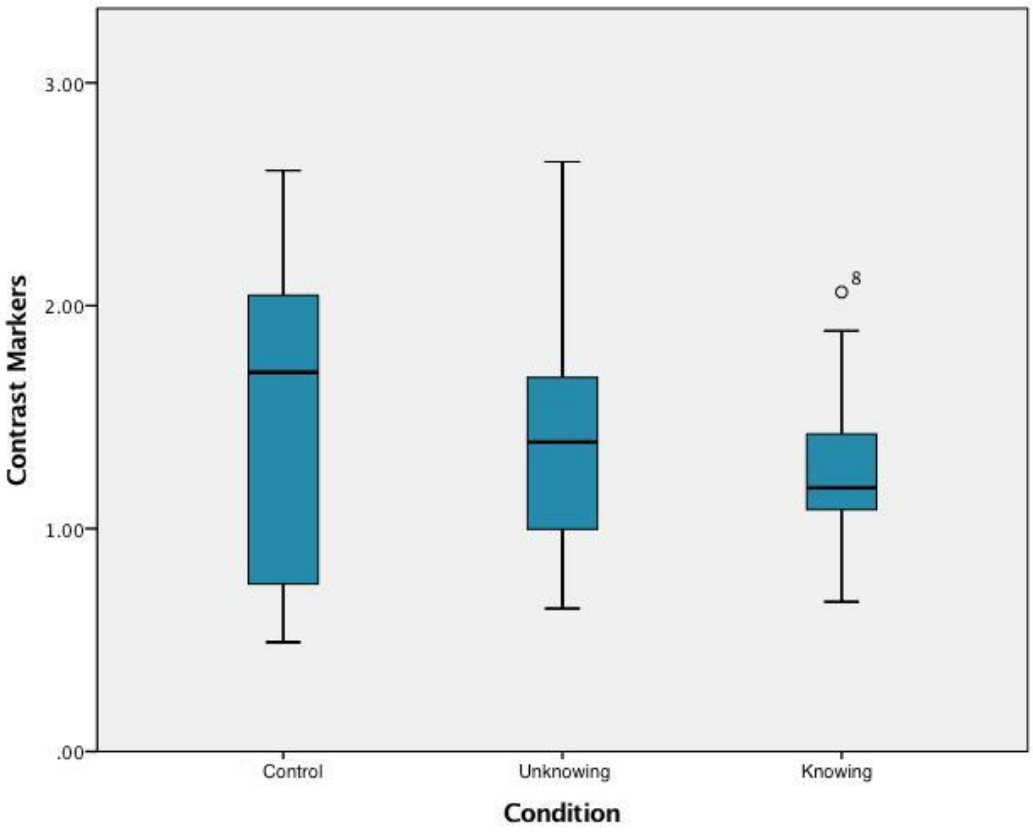

Figure 5.12 displays a boxplot showing the frequencies of contrast markers across conditions.

There is no significant difference in the frequency of contrast markers across conditions, as confirmed by a non-parametric Kruskal Wallis test ($H_{(2)}$ = 1.210, p=0.55). There is however, a significant difference in the frequency of negation across conditions. A non-parametric Kruskal Wallis test confirms an omnibus effect of condition on negation frequency ($H_{(2)}$ = 6.088, p<0.05), with a post hoc pairwise comparison with Dunn's test showing that there is significantly more negation in Unknowing than Control (p=0.04), but no significant difference between Unknowing and Knowing (p=0.535) nor Control and Knowing (p=0.79).

### 5.6.4 Deliberation Quality

Table 5.6 details the mean number of changes from a given stance position to another per participant over the course of the dialogue for each condition, as well as the total number of possible alternatives considered.

Fig. 5.12 Boxplot of Contrast Marker Frequencies



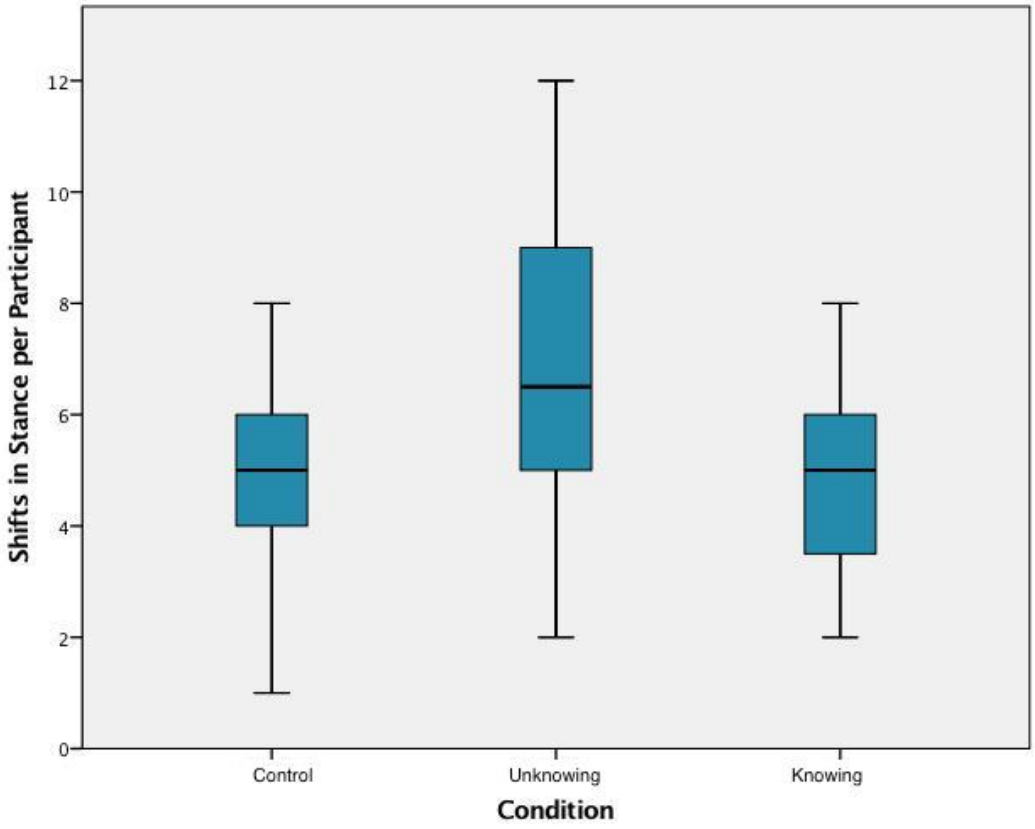

| Condition | Stance Shifts | Alternatives Considered | Point in Conversation First Opinion Reached | Point in Conversation Last Opinion Reached |
|---|---|---|---|---|
| Control | 4.85 | 3.10 | 14.50% | 56.70% |
| Std. Dev. | 1.843 | 0.968 | 11.10 | 27.77 |
| Unknowing | 6.80 | 3.30 | 13.36% | 67.89% |
| Std. Dev. | 2.628 | 0.657 | 12.63 | 17.64 |
| Knowing | 4.55 | 2.75 | 19.37% | 56.53% |
| Std. Dev. | 1.605 | 0.550 | 17.54 | 30.33 |
| Total Mean | 5.40 | 3.05 | 15.75% | 60.37% |
| Std. Dev. | 2.271 | 0.769 | 14.04 | 25.95 |

Table 5.6 Mean Stance shifts During Dialogue and Possible Solutions Considered per participant by condition

There are a third more stance shifts in the Unknowing condition than the Control and Knowing conditions. A Kruskal Wallis non-parametric test shows that there is a significant omnibus effect of condition on the number of stance shifts traversed by a

participant ($H_{(2)}$ = 9.559 p<0.01). A planned pairwise post hoc comparison using the Dunn's test shows that there are significantly more stance shifts in the Unknowing condition than the Knowing condition (p=0.01) but no confirmed significant effect between Unknowing and Control (p=0.06). Figure 5.13 displays a boxplot showing the mean number of shifts in stance position by participant across conditions.

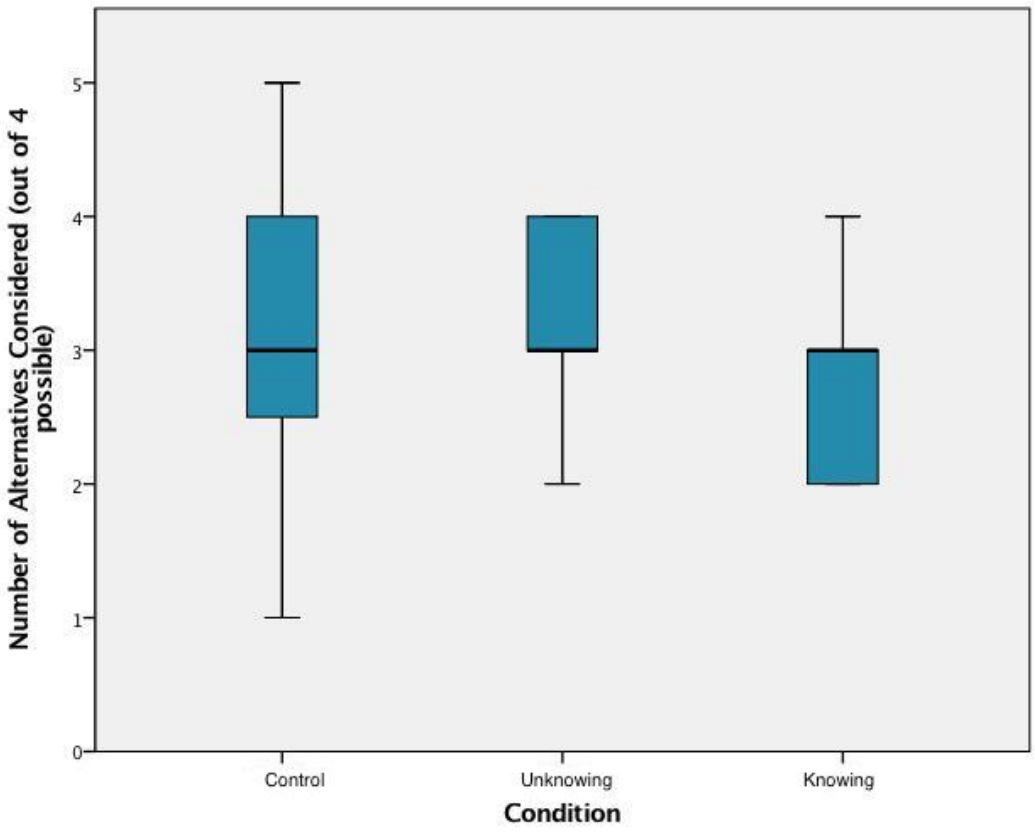
Fig. 5.13 Boxplot of Total Number of Stance Shifts per Participant




There is an omnibus effect of condition on number of possible solutions considered ($H_{(2)}$ = 6.146 p<0.05). There are more possible solutions considered in the Unknowing condition than the Knowing condition (p=0.04). However, there is no significant difference between Knowing and Control (p=0.33) and nor Control and Unknowing conditions (p=.1.00).

Figure 5.14 displays a boxplot showing the mean number of possible solutions considered per participant across conditions.

Fig. 5.14 Boxplot of Number of Solutions Considered




There is no significant difference in when the first opinion ($H_{(2)}$ = 1.429 p=0.49) nor last opinion ($H_{(2)}$ = 1.712 p=0.43) is reached across conditions.

### Amount of Conversation Spent in Concord versus Discord

Table 5.7 provides details of the mean percentage of turns in which participant A and B had matching and opposing stance states across conditions.

| Condition | Turns with Matching Stance | Turns with Opposing Stance |
|---|---|---|
| Control | 39.42% | 60.58% |
| Std. Dev. | | 0.14 |
| Unknowing | 48.27% | 51.73% |
| Std. Dev. | | 0.17 |
| Knowing | 32.74% | 67.26% |
| Std. Dev. | | 0.27 |
| Total Mean | 40.15% | 59.85% |
| Std. Dev. | | 0.20 |

Table 5.7 Mean percent of dialogue in which participant A and B had matching and opposing stance positions

The distributions show approximately 16% difference in the ratio of opposing and matching stance positions between Knowing and Unknowing conditions, with more turns covered with opposing stances in the Knowing condition and more matching stances in the Unknowing Condition. However, a non-parametric Kruskal Wallis test find no significant effect of condition on the distribution of oppositional and matching stance states amongst participants ($H_{(2)}$ = 3.850 p=0.15).

### 5.6.5 Pronoun Usage

Table 5.8 provides mean frequencies of self-referencing, other-referencing and collective pronoun frequencies, per 100 word,per dyad, for each condition.

| Condition | | Me/My/I | You/Yours | We/Ours | Total Pronouns |
|---|---|---|---|---|---|
| Control | Mean | 1.30 | 0.47 | 0.64 | 36.10 |
| | Std. Dev. | 0.51 | 0.26 | 0.43 | 9.92 |
| Unknowing | Mean | 1.44 | 0.49 | 0.74 | 50.10 |
| | Std. Dev. | 0.51 | 0.25 | 0.33 | 19.82 |
| Knowing | Mean | 1.16 | 0.61 | 0.91 | 44.40 |
| | Std. Dev. | 0.38 | 0.32 | 0.46 | 14.41 |
| Total | Mean | 1.30 | 0.52 | 0.76 | 43.53 |
| | Std. Dev. | 0.47 | 0.27 | 0.41 | 15.85 |

Table 5.8 Individual, Other and Collective Personal Pronouns Normalised per 100 words and Mean frequencies of Total Pronouns per dyad

A non-parametric Kruskal Wallis test shows no effect of condition on the frequency of self-referencing personal pronouns per 100 ($H_{(2)}$ = 1.987, p=0.37), nor other-referencing personal pronouns ($H_{(2)}$ = 1.443, p=0.49). There is also no significant effect of condition on the frequency of collective personal pronouns per 100 as confirmed by a non-parametric Kruskal Wallis test ($H_{(2)}$ = 1.215, p=0.55).

### 5.6.6 Post-hoc Analysis

Some additional post-hoc analysis was conducted on the experiment data. As a follow on from the previous experiment, to assess what impact our stance marking insertions had upon the levels of agreement and disagreement within the dialogues, the frequencies of turn-initial Agreement, Disagreement and Update markers, were examined.

#### Turn-Initial Agreement, Disagreement and Update Markers

Table 5.9 details the mean frequencies of turn-initial Agreement, Disagreement and Update Markers per 100 turns, per condition:

| Condition | | TI Agreement | TI Disagreement | TI Update |
|---|---|---|---|---|
| Control | Mean | 25.15 | 22.70 | 2.10 |
| | Std. Deviation | 18.51 | 30.84 | 1.73 |
| Unknowing | Mean | 22.25 | 10.18 | 2.0 |
| | Std. Deviation | 14.45 | 3.79 | 1.08 |
| Knowing | Mean | 20.84 | 12.20 | 4.70 |
| | Std. Deviation | 19.07 | 5.08 | 1.68 |
| Total | Mean | 27.90 | 16.37 | 3.47 |
| | Std. Deviation | 31.62 | 11.17 | 5.61 |

Table 5.9 Turn-Initial Markers

Figure 5.15 displays a boxplot for turn-initial agreement marker frequencies.

Fig. 5.15 Boxplot of Turn-initial Agreement Marker Frequencies

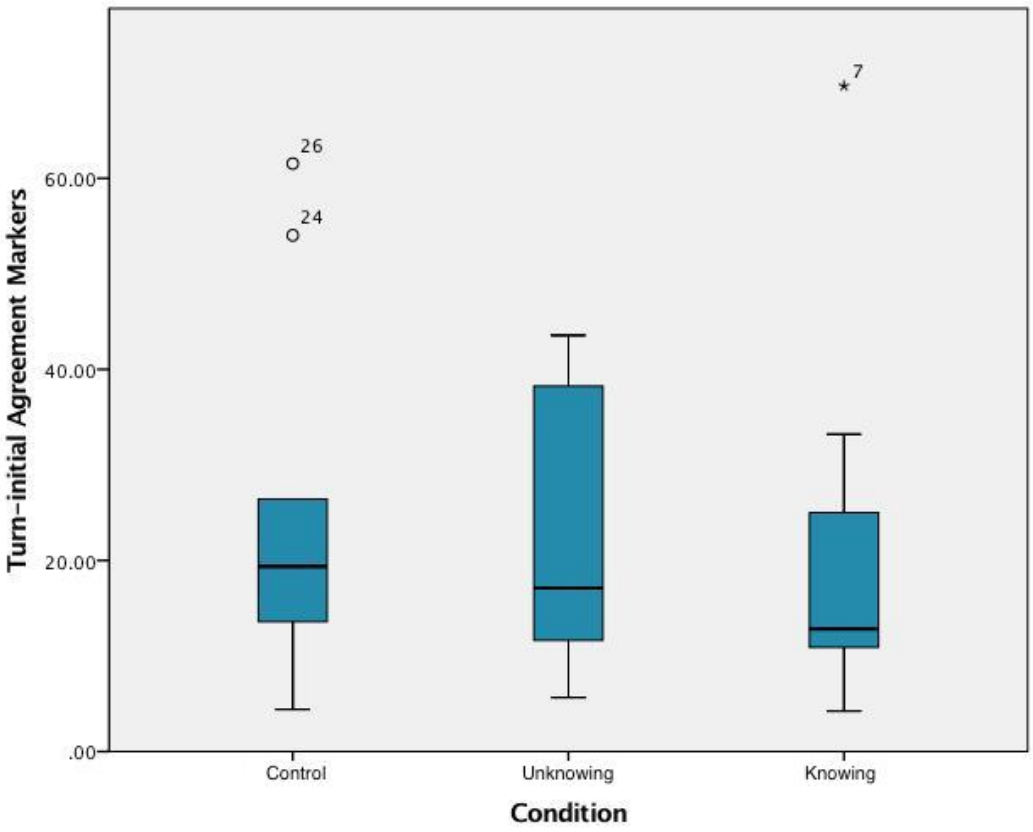



Figure 5.16 displays a boxplot for turn-initial disagreement marker frequencies.

Fig. 5.16 Boxplot of Turn-initial Disagreement Marker Frequencies

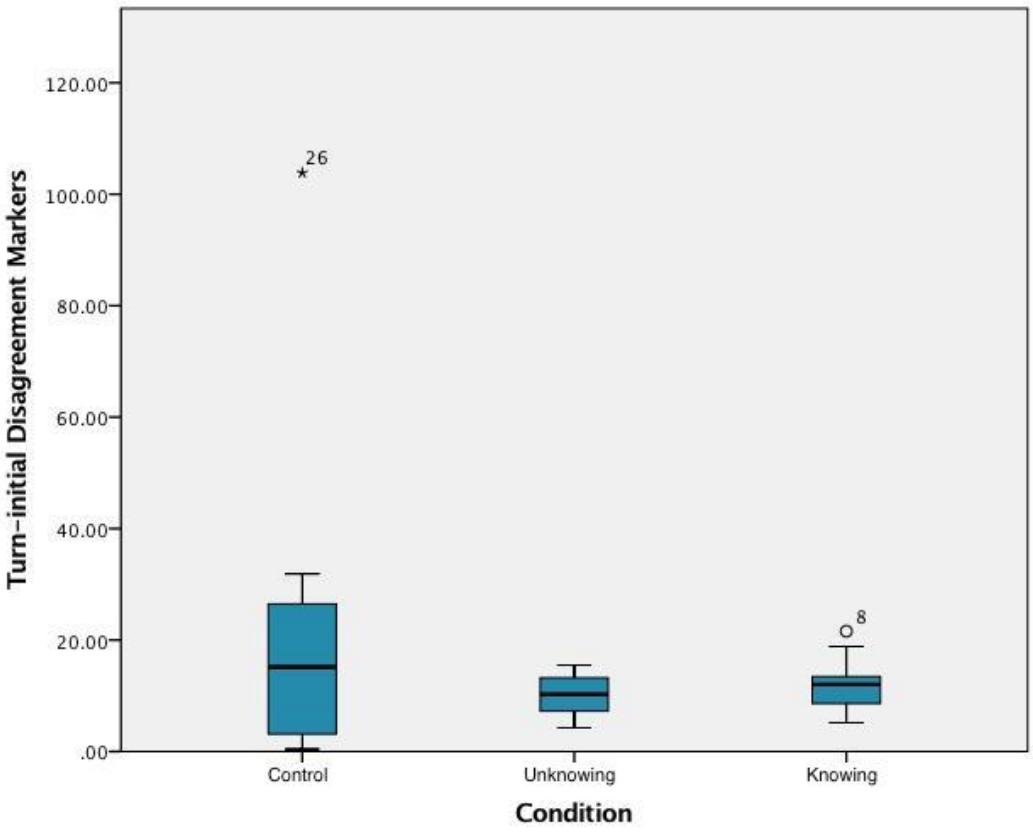
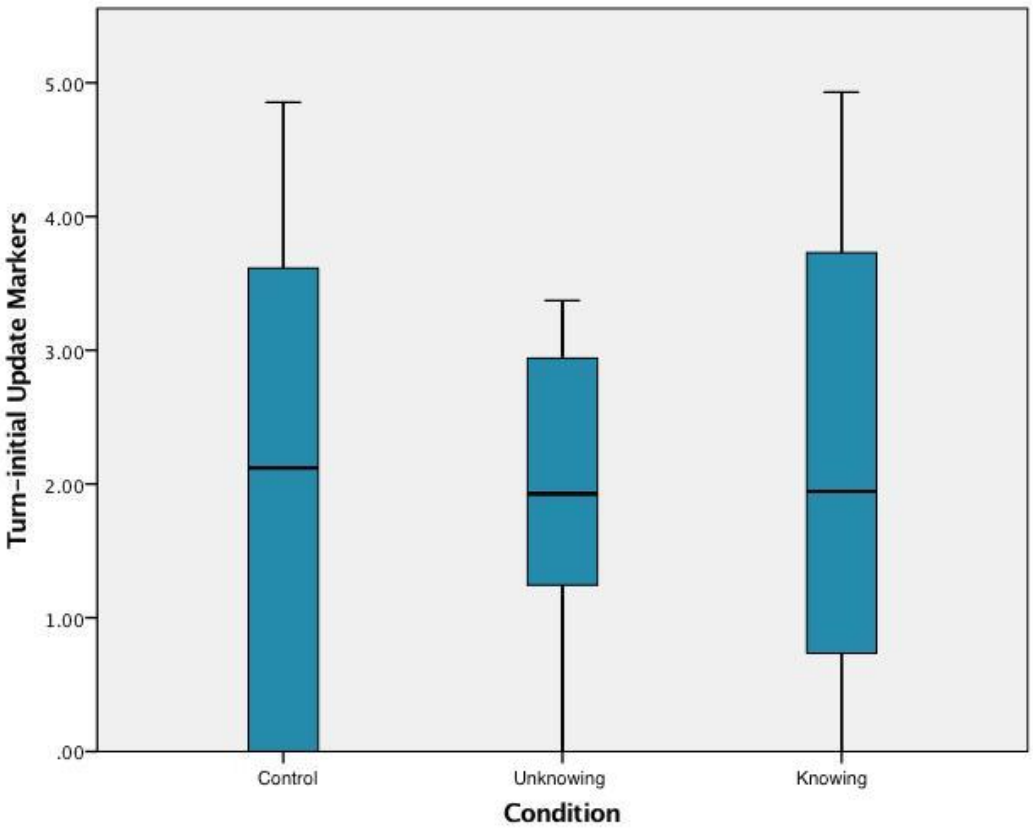



Figure 5.17 displays a boxplot for turn-initial update marker frequencies.

Fig. 5.17 Boxplot of Turn-initial Update Marker Frequencies




A non-parametric Kruskal Wallis test shows no statistically significant effect of condition on the frequency of turn-initial Agreement Markers, ($H_{(2)}$ = 0.794, p= 0.67), turn-initial Disagreement Markers, ($H_{(2)}$ = 0.560, p= 0.76), nor turn-initial update markers ($H_{(2)}$ = 0.148, p= 0.93).

## 5.7 Discussion

### 5.7.1 Stance Shifts and Alternatives considered

In line with our prediction, the results show that framing statements as unknowing led to more deliberation in the dialogues. Not only were there a higher numbers of shifts in stance, indicating a thorough deliberation going back and forth over the possible solutions, there was also a fuller exploration of the total possible solutions (i.e. participants in the Unknowing condition were more likely to consider all of the four possible outcomes, and consider each person to be ejected rather than just sticking to a one or two). This supports our hypothesis that 'I think' positioned assertions as more 'knowing', and caused a closing down of the discussion and reduced the quality and coverage of the deliberation. Conversely, as predicted, presenting opinions as

'unknowing' improved the deliberative quality of the dialogue, with more ideas and positions exchanged and elaborated on.

### 5.7.2 Message Construction: typing speed, time and self edits

The results show that the introduction of the knowing stance marker 'I think' leads to fewer words per turn, i.e. shorter, or more terse responses. This greater efficiency in the construction of dialogue turns, suggests that the introduction of the knowing stance marker leads to more direct exchange of opinions. This directness is also supported by the fewer edits during turn construction in this condition. Less care is taken in the Knowing condition to alter the message prior to relaying it to a conversational partner, perhaps leading to less delicately constructed or polite turns, but more direct and less guarded opinion exchange. The results show that prefacing statements with a knowing preface (i.e. 'I think') forecloses the conversation.

### 5.7.3 Speaker Commitment

Counter to our predictions there was no significant effect of condition on the frequency of expressions of uncertainty. However, significantly more certainty adverbials are employed by participants in the Unknowing condition compared to the Control condition. This suggests that framing contributions as unknowing creates an environment in which participants are more likely to make manifest their commitment to a stance by upgrading the epistemic strength of a statement through certainty adverbials; as solutions are discussed more and potentially co-constructed, once a stance is established it can be committed to with greater conviction by participants in the Unknowing Condition. So, although the Knowing Condition features less guarded and more direct messages as indicated in the manner in which they are constructed, it is in the Unknowing Condition that speakers commit more firmly to the substantive essence of their utterance.

Interpreting these results together suggests that the introduction of 'Do you think' opens up the dialogue, inviting further elaboration of the topic at hand, while introducing 'I think' closes down the dialogue and limits the deliberative quality of the discussion. 'Do you think' positions the speaker in a position of unknowing epistemic status, and also directly invokes the hearer to collaborate in the co-construction of a joint stance. In the Unknowing Condition, stance positions are more explicitly emphasised through certainty adverbials and negation, i.e. when something is important, speakers take care to make clear the focus of their stance and emphasise the strength of commitment to a given proposition. In part this may be due to the fact that 'do you think' directly invites input and therefore greater care is taken to make clear

exactly what the opinion to which they are attaching themselves is. The interactive negotiation of the stance is more exaggerated. Conversely, the introduction of 'I think' to the dialogue has the opposite effect: the presentation of a knowing stance, leads to less consideration and more conviction among participants, demonstrated through fewer edits when constructing responses and more terse and direct turns. Opinions are expressed plainly and without additional specification.

### 5.7.4 Pronoun Usage

Counter to our hypothesis, manipulations to the degree of knowingness did not affect the use of personal pronouns. This may indicate that framing contributions as unknowing doesn't necessarily lead to more joint construction of stance, although the fact that individuals spent more time considering a shared stance poses a challenge to this. Or it may be that personal pronouns do not provide a robust enough indicator of joint versus individual stance construction in this context.

## 5.8 Conclusion

In this Chapter the causal effects of epistemic status, as expressed through particular stance markers, on the deliberative quality of a dialogue were investigated using an experimental approach. Framing a statement as knowing has a significant impact on the deliberative quality of a dialogue and decreases the likelihood that participants will consider multiple possible solutions, shifting their opinion fewer times before reaching a concluding stance. Furthermore, participants in the Unknowing Condition, spent a larger proportion of dialogues considering one another's stance. This suggests that, within a discussion dialogue, the framing of a statement in a unknowing way can lead to a more flexible deliberation process and a greater willingness to engage with alternative viewpoints. Furthermore, being more considerate of one another's views was not necessarily to the detriment of expressing a position with conviction; one explanation for the higher frequencies of certainty adverbials could be that adding 'do you think' led to greater displays of speaker commitment to a stance one a stance position was settled upon. However, although the unknowing condition differed from the control, it did not differ from the knowing condition to a level deemed significant ($p= 0.08$), which makes this proposal less concrete and further investigation would be needed to confirm if this is the case.

Framing a statement as knowing affects the ways in which individuals produce messages; specifically, framing as messages as knowing led to fewer insert-edits. This suggests that there is less care taken in the construction of messages, and less conscientious effort put into producing polite, or considered turns.

The length of message was affected by condition but it is not possible to determine whether framing as contributions as knowing led to shorter responses or framing as contributions as unknowing led to longer responses, as neither knowing nor unknowing differed from the control). However, it does appear that the two markers are having acting upon different ends of a spectrum in terms of opening up or closing down the dialogue.

Shorter messages are typically more direct and the fewer insert-edits may reflect decreased guardedness. one way to interpret this is that by prefacing statements with 'I think', the context is set for the exchange of opinions; by introducing a stance with a knowing marker, the appropriateness of a response that is equally direct is established. Conversely, introducing a contribution with 'Do you think' simply puts it on the table for further discussion, directly inviting for input, and therefore a lengthier response.

Overall it seems that marking stances with a knowing preface leads to more direct and unguarded exchanges, but does not improve the deliberative quality of the dialogues. Conversely, prefacing statements with the unknowing preface 'do you think' encourages a more collaborative deliberation, in which more possible solutions are considered in turn before a final decision is reached.

# Chapter 6

# Discussion and Conclusion

In this thesis we have presented an experimental approach for examining the causal effects of exposed (dis)agreement and stance constructions with differing levels of knowingness in dialogue. In particular it has addressed how such constructions can serve to open up or close down discussion dialogues. This has focused on a comparison of exposed agreement and disagreement with an interlocutor and the framing of an opinion as knowing or unknowing. In both experiments a motivating factor was a desire to better understand how the construction and framing of discursive content could support constructive engagement and lead to more considered discussion dialogues, with additional viewpoints considered.

## 6.1 Introduction

The results of this thesis suggest that the ways in which people negotiate the various positions proposed during an interaction is managed through a range of subtle linguistic devices, and that the consequences of how a speaker presents their position can have non-trivial effects. In line with the literature, we demonstrate that constructions which explicitly challenge or dismiss another speaker's position (e.g. you're wrong), rarely occur, and when artificially inserted they disrupt dialogues and frustrate the potential for deliberation. These findings concur with Angouri (2012); Chiu (2008), which highlight that impoliteness is a key influential factor that can affect how speaker challenges are interpreted and responded to. Thus, despite the intuition that adversity is a necessary precursor to considered and thoughtful dialogue, the ways in which speakers convey opposition is essential to the constructive outcomes of an interaction.

However, although explicitly announcing a disagreement is rare, positioning content as oppositional is still possible and is achieved through in a variety of ways. Furthermore, it is not simply about explicitly marking what is said as oppositional,

but rather that it is newsworthy, substantive or fundamental to the integration and progression of the ongoing dialogue. While directly disagreeing with an interlocutor can lead to disengagement which closes down the dialogue space, more subtle devices which mark what is being said as a substantive, or with potential consequences for the ongoing interaction can play an important role in deliberative dialogues. This interpretation supports the wider work on discourse markers and pragmatic particles, which have been shown to aid hearer's ability to recall aspects of a dialogue (Fox Tree and Schrock, 1999, 2002; Liu and Tree, 2012) – and mark that attention ought to be paid to what is being said.

The other factor at play, is the management of interpersonal relations. Qualitative studies have revealed specific phenomena associated with disagreement, such as delays and the prefacing of disagreement with agreement (for example, "Oh yes, she is a lovely girl, but she can be problematic") (Pomerantz, 1987); this attends to something other than marking content as substantive, namely the face of participants in an interaction. Heritage's work on epistemic status highlights how this continually shifting engine of balance and imbalance drives contributions in a dialogue. The experimental work on knowingness presented in this thesis highlights the significance of how a speaker's commitment to a positions can directly influence the chance that alternative solutions are considered. The implications of this suggest that how people put forth a position, and the perceived attachment they have to it, alters co-conversants ability to comment on, and collaborative co-construct a position. If issued with less overt commitment there is additional space for negotiation of the specific terms of agreement or consensus. How propositions are put forth, therefore, can be a significant in the creation of a dialogical context in which deliberation can occur.

## 6.2 Summary of Contribution

Chapter 3 presented a study of a large corpus of everyday conversations, in which the distribution of speaker's markers of (dis)agreement, updates, contrast and emphasis was compared across samples of direct speech and reports of their own speech and the speech of others. The analysis showed that exposed disagreement occurs very rarely, but that there are a variety of other ways in which speaker's signal that their position is contrasting to what has come before. One of the most interesting among these is the way people mark their rights to speak about something. The use of such resources as reported speech and prefacing incongruent content with discourse markers (e.g. 'well') can be important when constructing and presenting a stance position and in the management epistemic status and face concerns.

In line with existing research, the corpus analysis of ordinary conversation highlighted that exposed and direct disagreement occurs very rarely. However, exposed agreement also occurred relatively infrequently. The context of reported speech was investigated as a site in which the potential face-threat of disagreeing with an interlocutor was removed or reduced. While exposed disagreement was not more common in reported speech, it was distinct from direct speech (non-reports), in a number of ways, which support the interpretation of reported speech as an interactive evidential proposed by Clift (2006a,b), and its role as a stance marker. The content framed as reports highlighted the substance of the difference, the oppositional stance, as evident by higher frequencies of contrastive markers and negation. This highlighted the various ways in which speakers mark oppositional content and employ evidential devices to convey their position to their co-conversants.

What the corpus study could not shed light on was what happens when people do directly disagree with a co-conversant in an unmitigated manner. While Chiu (2008) found that polite disagreement was beneficial to advancing dialogues and led to novel contributions, what of impolite or more direct ways of exposing disagreement?

An experiment was presented that used a method that allowed fine-grained manipulations of text based dialogues in real-time. Exposed agreement and disagreement fragments were inserted into a discussion dialogue. As exposed disagreement occurs so rarely in natural conversation this afforded us a unique opportunity to ascertain the interactional effect of exposed disagreement on a dialogue. Furthermore, Chiu (2008) suggests that agreement can be problematic for discursive contexts. The effect of agreement on the expansion or contraction of dialogic space has garnered much less attention, and in constructing our experiment in chapter 4, it was possible for us to directly compare the effect of agreement and disagreement on the course of a dialogue.

The findings showed that the insertion of explicit disagreement violates the conventions of polite dialogue, leading participants to put more effort into the production of their replies. Insertions of exposed disagreement disrupt dialogues, bringing the topic of disagreement directly into the conversation, provoking clarification requests and resulting in a greater number of self-edits when formulating turns. The insertion of disagreement also led to more instances of exposed agreement, suggesting that dialogue partners co-operate to redress the face-threat of disagreement. Conversely, exposed agreement insertions were not as incongruous and had less disruptive impact on the ensuing dialogues. However, introducing agreement into the dialogue did lead to greater deliberation, with more alternative scenarios considered by participants during the task. One interpretation of this would be to conclude that being polite, and thus creating a supportive discussion environment is more important that exposing contrary viewpoints.

Contrary to the findings on polite disagreement by Chiu (2008), direct disagreement did not improve the quality of deliberation. Rather, in line with Pomerantz (1984a), chapter 4 demonstrated the socially problematic effect of disagreeing in a direct and unmitigated manner with an interlocutor. The insertions of explicit disagreement clearly violated norms and conventions of polite dialogue, leading to extra composition and editing of reply turns in response. Furthermore, while in part, it was hypothesised that disagreement could be productive and serve to move the dialogue on, direct disagreement did not have this effect: the dialogue space was closed down by the introduction of exposed disagreement; thus it was not interpreted as, nor did it foster, constructive engagement. By comparison, however, the introduction of agreement fragments into a dialogue did lead to more possible viewpoints being consulted.

Chapter 5 facilitated the examination of a more subtle facet of discursive dialogue, namely the framing of a stance as knowing or unknowing. This enabled the manipulation of oppositional content in a less face-threatening manner, but with differing degrees of 'knowingness'. Rather than intervening in the presentation of direct agreement or disagreement, in this experiment we were able to manipulate the way in which propositional content was presented to interlocutors. Building on the work of Heritage (2012a,b), we were able to systematically test the precise effects of altering the knowingness of speaker contributions, and demonstrate a causal relation to the deliberative quality of a dialogue.

The results show that how knowingly a position is presented can have significant effects on the quality of the ensuing deliberation. In particular, when assertions were framed as unknowing and directly invited participants to co-construct stance positions, more diverse opinions were consulted by both parties in the dialogue. Marking speaker stance as knowing led to more direct and unguarded exchanges, but did not improve the deliberative quality of the dialogues. Conversely, framing statements as less knowing increased the likelihood that participants engage with and consider more alternative viewpoints before reaching a concluding stance, thus increasing the deliberative quality of a dialogue. By investigating the impact of altering the epistemic framing of a stance as 'knowing' or 'unknowing', this thesis demonstrates the interactional effect of more subtle devices for presenting oppositional content in dialogue.

In combination, these studies have advanced the understanding of the causal effects of disagreement and degrees of knowingness in discursive contexts, namely that while exposed disagreement is problematic and does not foster constructive engagement, reducing the degree of knowingness with which assertions are put forth

can contribute to a discussion context in which the reconsideration of multiple stance positions are encountered.

## 6.3 Potential Applications of the Study

By investigating the impact of altering the epistemic framing of a stance as 'knowing' or 'unknowing', this thesis assesses the interactional effect of more subtle devices for presenting oppositional content in dialogue. Potential applications of this research include training guidelines for facilitators of deliberation or individuals developing tools which aim to promote more considered debate. There is also a wider implication for computational approaches to argumentation, in that it highlights the importance of interpersonal dynamics and identifies stance construction as a key resource for formulating polite arguments; consequently, stance markers may be particularly useful for detecting informal argument in socially generated data.

## 6.4 Limitations and Further Work

The experimental work presented in this thesis focuses on dyadic interaction. How epistemic status and the management of face is negotiated in multi-party interaction would need further consideration. Additionally, the specific ways in which the findings of the thesis, such as how constructive engagement in the deliberative process can be improved by presenting contributions as less knowing could actually be integrated into online tools would need additional research.

Recent developments to the DiET Toolkit have made it possible to log the character by character insertions and deletions made while constructing a turn to relay to a interlocutor. Since both of our experiments show that an effect of intervention on the message construction, formulation and editing before sending a turn, it a series of experiments using this additional functionality could provide useful insights into the cognitive processes involved in managing the reception and response to socially problematic content in contexts in which the management of diverse viewpoints is critical.

# References


Rob Abbott, Marilyn Walker, Pranav Anand, Jean E Fox Tree, Robeson Bowmani, and Joseph King. How can you say such things?!?: Recognizing disagreement in informal political argument. In *Proceedings of the Workshop on Languages in Social Media*, pages 2–11. Association for Computational Linguistics, 2011.

Karin Aijmer. I think—an english modal particle. *Modality in Germanic languages: Historical and comparative perspectives*, 99:1, 1997.

Karin Aijmer. *English discourse particles: Evidence from a corpus*, volume 10. John Benjamins Publishing, 2002.

A.Y. Aikhenvald. *Evidentiality*. Oxford linguistics. Oxford University Press, 2006.

Francisco Alonso-Almeida and Laura Cruz-García. The value of may as an evidential and epistemic marker in english medical abstracts. *Studia anglica posnaniensia*, 46(3):59–73, 2011.

Pranav Anand, Marilyn Walker, Rob Abbott, Jean E. Fox Tree, Robeson Bowmani, and Michael Minor. Cats rule and dogs drool!: Classifying stance in online debate. In *Proceedings of the 2nd workshop on computational approaches to subjectivity and sentiment analysis*, pages 1–9. Association for Computational Linguistics, 2011.

Jo Angouri. Managing disagreement in problem solving meeting talk. *Journal of Pragmatics*, 44(12):1565–1579, 2012.

Jo Angouri and Miriam A Locher. Theorising disagreement. *Journal of Pragmatics*, 44(12):1549–1553, 2012.

Jo Angouri and Theodora Tseliga. "you have no idea what you are talking about!" from e-disagreement to e-impoliteness in two online fora. *Journal of Politeness Research. Language, Behaviour, Culture*, 6(1):57–82, 2010.

Paul Atkinson. Medical discourse, evidentiality and the construction of professional responsibility. *Talk, work and institutional order: Discourse in medical, mediation and management settings*, 75108, 1999.

Farah Benamara, Baptiste Chardon, Yannick Mathieu, Vladimir Popescu, and Nicholas Asher. How do negation and modality impact on opinions? In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, pages 10–18. Association for Computational Linguistics, 2012.

Sabine Bergler. Conveying attitude with reported speech. In *Computing attitude and affect in text: Theory and applications*, pages 11–22. Springer, 2006.

Douglas Biber. Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2):97–116, 2006.

Douglas Biber and Edward Finegan. Adverbial stance types in english. *Discourse processes*, 11(1):1–34, 1988.

Douglas Biber and Edward Finegan. Styles of stance in english: Lexical and grammatical marking of evidentiality and affect. *Text-interdisciplinary journal for the study of discourse*, 9(1):93–124, 1989.

Douglas Biber, Stig Johansson, Geoffrey Leech, Susan Conrad, Edward Finegan, and Randolph Quirk. *Longman grammar of spoken and written English*, volume 2. MIT Press, 1999.

J. Bohman and William Rehg. *Deliberative Democracy: Essays on Reason and Politics*. Overtures to Biblical Theology. MIT Press, 1997.

Susan E Brennan and Justina O Ohaeri. Why do electronic conversations seem less polite? the costs and benefits of hedging. In *ACM SIGSOFT Software Engineering Notes*, volume 24, pages 227–235. ACM, 1999.

P. Brown and S.C. Levinson. *Politeness: Some Universals in Language Usage*. Studies in Interactional Sociolinguistics. Cambridge University Press, 1987.

Lou Burnard. *British National Corpus: Users Reference Guide British National Corpus Version 1.0*. Oxford Univ. Computing Service, 1995.

Wallace Chafe. Evidentiality in english conversation and academic writing. *Evidentiality: The linguistic coding of epistemology*, 20:261–273, 1986.

Massimo Chindamo, Jens Allwood, and Elisabeth Ahlsen. Some suggestions for the study of stance in communication. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*, pages 617–622. IEEE, 2012.

Ming Ming Chiu. Flowing toward correct contributions during group problem solving: A statistical discourse analysis. *the journal of the learning sciences*, 17 (3):415–463, 2008.

Herbert H Clark and Susan E Brennan. Grounding in communication. *Perspectives on socially shared cognition*, 13(1991):127–149, 1991.

Herbert H Clark and Richard J Gerrig. Quotations as demonstrations. *Language*, pages 764–805, 1990.

R. Clift. Getting there first: non-narrative reported speech in interaction. In E. Holt and R. Clift, editors, *Reporting Talk: Reported Speech in Interaction*, Studies in Interactional Sociolinguistics, chapter 5, pages 120–149. Cambridge University Press, 2007.

Rebecca Clift. Indexing stance: Reported speech as an interactional evidential1. *Journal of sociolinguistics*, 10(5):569–595, 2006a.

Rebecca Clift. Indexing stance: Reported speech as an interactional evidential1. *Journal of sociolinguistics*, 10(5):569–595, 2006b.

J. Coates. *The Semantics of the Modal Auxiliaries*. Croom Helm linguistics series. Croom Helm, 1983.

Marcus Colman, Patrick GT Healey, et al. The distribution of repair in dialogue. In *Proceedings of the 33rd annual meeting of the cognitive science society*, pages 1563–1568, 2011.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Taking a stance: a corpus study of reported speech. In *Proceedings of the 19th SemDial Workshop on the Semantics and Pragmatics of Dialogue (GoDial)*, Semdial Workshop, 2015a.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Shifting opinions: An experiment on agreement and disagreement in dialogue. In *Proceedings of the 19th SemDial Workshop on the Semantics and Pragmatics of Dialogue (GoDial)*, Semdial Workshop, 2015b.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. How natural is argument in natural dialogue? In *16th Workshop on Computational Models of Natural Argument*, New York, July 2016.

Shauna Concannon, Patrick GT Healey, and Matthew Purver. Opening up and closing down discussion: Experimenting with stance in conversation. In preparation.

Susan Conrad and Douglas Biber. Adverbial marking of stance in speech and writing. In *Evaluation in text: Authorial stance and the construction of discourse*, pages 56–73. Oxford University Press Oxford, 2000.

Elizabeth Couper-Kuhlen. Assessing and accounting. In E. Holt and R. Clift, editors, *Reporting Talk: Reported Speech in Interaction*, Studies in Interactional Sociolinguistics, chapter 4, pages 81–119. Cambridge University Press, 2007.

Mary J Culnan and M Lynne Markus. Information technologies. 1987.

Jonathan Culpeper. *Impoliteness*. Wiley Online Library, 2011.

Bronwyn Davies and Rom Harré. Positioning: The discursive production of selves. *Journal for the theory of social behaviour*, 20(1):43–63, 1990.

John Dewey. *Democracy and education*. Courier Corporation, 2004.

Angela Downing. "surely you knew!": Surely as a marker of evidentiality and stance. *Functions of Language*, 8(2):251–282, 2001a.

Angela Downing. "surely you knew!": Surely as a marker of evidentiality and stance. *Functions of Language*, 8(2):251–282, 2001b.

Paul Drew. What drives sequences? *Research on Language & Social Interaction*, 45 (1):61–68, 2012.

John W Du Bois. The stance triangle. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, pages 139–182, 2007.

Robert Englebretson. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, volume 164. John Benjamins Publishing, 2007.

Roy E Feldman. Response to compatriot and foreigner who seek assistance. *Journal of Personality and Social Psychology*, 10(3):202, 1968.

Kathleen Ferrara, Hans Brunner, and Greg Whittemore. Interactive written discourse as an emergent register. *Written communication*, 8(1):8–34, 1991.

Margaret Field. The role of factive predicates in the indexicalization of stance: A discourse perspective. *Journal of Pragmatics*, 27(6):799–814, 1997.

Susan Fitzmaurice. Subjectivity, intersubjectivity and the historical construction of interlocutor stance: from stance markers to discourse markers. *Discourse Studies*, 6(4):427–448, 2004.

Barbara A Fox. Evidentiality: Authority, responsibility, and entitlement in english conversation. *Journal of Linguistic Anthropology*, 11(2):167–192, 2001.

Jean E. Fox Tree. Listening in on monologues and dialogues. *Discourse processes*, 27(1):35–53, 1999.

Jean E. Fox Tree. Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34(1):37–55, 2002.

Jean E. Fox Tree and Josef C. Schrock. Discourse markers in spontaneous speech: Oh what a difference an oh makes. *Journal of Memory and Language*, 40(2): 280–295, 1999.

Jean E. Fox Tree and Josef C. Schrock. Basic meanings of< i> you know</i> and< i> i mean</i>. *Journal of Pragmatics*, 34(6):727–747, 2002.

Bruce Fraser. Pragmatic markers. *Pragmatics*, 6:167–190, 1996.

Rod Gardner. *When listeners talk: Response tokens and listener stance*, volume 92. John Benjamins Publishing, 2001.

Alexandra Georgakopoulou. Arguing about the future: On indirect disagreements in conversations. *Journal of Pragmatics*, 33(12):1881–1900, 2001.

Howard Giles and Philip Smith. *Accommodation theory: Optimal levels of convergence*. 1979.

Howard Giles, Justine Coupland, and Nikolas Coupland. *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press, 1991.

Dylan Glynn. Subjectivity and epistemicity an annotated bibliography. *Subjectivity and Epistemicity*, page 349.

Erving Goffman. On face-work. *Interaction ritual*, pages 5–45, 1967a.

Erving Goffman. Interaction ritual: essays on face-to-face interaction. 1967b.

Erving Goffman. *Forms of talk*. University of Pennsylvania Press, 1981.

Charles Goodwin. The interactive construction of a sentence in natural conversation. *Everyday language: Studies in ethnomethodology*, pages 97–121, 1979.

Charles Goodwin. *Conversational organization: Interaction between speakers and hearers*. Academic Press, 1981.

Charles Goodwin and Marjorie Harness Goodwin. 6 assessments and the construction of context. *Rethinking context: Language as an interactive phenomenon*, (11): 147, 1992.

M. Haakana. Reported thought in complaint stories. In E. Holt and R. Clift, editors, *Reporting Talk: Reported Speech in Interaction*, Studies in Interactional Sociolinguistics, chapter 6, pages 150–178. Cambridge University Press, 2007.

Jürgen Habermas. The theory of communicative action, volume i. *Boston: Beacon*, 1984.

Pentti Haddington. Stance taking in news interviews. *SKY Journal of Linguistics*, 17:101–142, 2004.

Pentti Haddington. *The intersubjectivity of stance taking in talk-in-interaction*. University Press, 2005.

Mia Halonen. Person reference as a device for constructing experiences as typical in group therapy. *Conversation analysis and psychotherapy/ed. by Anssi Peräkylä*, 2008.

Kazi Saidul Hasan and Vincent Ng. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *EMNLP*, pages 751–762, 2014.

Patrick Healey. Interactive misalignment: The role of repair in the development of group sub-languages. *Language in Flux. College Publications*, 212, 2008.

Patrick Healey, Matthew Purver, James King, Jonathan Ginzburg, and Greg Mills. Experimenting with clarification in dialogue. In *Proceedings of the 25th annual meeting of the cognitive science society*, pages 539–544, 2003.

John Heritage. Epistemics, conversation analysis, and'post-analytic'ethnomethodology: A rebuttal.

John Heritage. A change-of-state token and aspects of its sequential placement. *Structures of social action: Studies in conversation analysis*, pages 299–345, 1984.

John Heritage. Oh-prefaced responses to inquiry. *Language in society*, 27(3): 291–334, 1998.

John Heritage. The limits of questioning: Negative interrogatives and hostile question content. *Journal of pragmatics*, 34(10):1427–1446, 2002.

John Heritage. Conversation analysis as social theory. *The new Blackwell companion to social theory*, pages 300–320, 2008.

John Heritage. The epistemic engine: Sequence organization and territories of knowledge. *Research on Language & Social Interaction*, 45(1):30–52, 2012a.

John Heritage. Epistemics in action: Action formation and territories of knowledge. *Research on Language & Social Interaction*, 45(1):1–29, 2012b.

John Heritage and Steven Clayman. *Talk in action: Interactions, identities, and institutions*. Wiley. com, 2010.

John Heritage and Geoffrey Raymond. The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly*, 68(1):15–38, 2005.

Janet Holmes. Functions of you know in women's and men's speech. *Language in society*, 15(01):1–21, 1986.

Janet Holmes. Hedges and boosters in women's and men's speech. *Language & Communication*, 10(3):185–205, 1990.

E. Holt. 'i'm eyeing your chop up mind': reporting and enacting. In E. Holt and R. Clift, editors, *Reporting Talk: Reported Speech in Interaction*, Studies in Interactional Sociolinguistics, chapter 3, pages 47–80. Cambridge University Press, 2007.

E. Holt and R. Clift, editors. *Reporting Talk: Reported Speech in Interaction*. Studies in Interactional Sociolinguistics. Cambridge University Press, 2007.

Elizabeth Holt. Reporting and reacting: Concurrent responses to reported speech. *Research on Language and Social Interaction*, 33(4):425–454, 2000.

Thomas Holtgraves. Yes, but... positive politeness in conversation arguments. *Journal of Language and Social Psychology*, 16(2):222–239, 1997.

L. Hoye. *Adverbs and Modality in English*. English language series. Longman, 1997.

Susan Hunston. Conflict and consensus: Construing opposition in applied linguistics. *Strategies in academic discourse. Amsterdam: Benjamins*, pages 1–15, 2005.

Susan Hunston. Using a corpus to investigate stance quantitatively and qualitatively. *Stancetaking in discourse*, pages 27–48, 2007.

Susan Hunston and Geoffrey Thompson. *Evaluation in Text: Authorial Stance and the Construction of Discourse: Authorial Stance and the Construction of Discourse*. Oxford University Press, 2000.

Susan Hunston et al. *Corpus approaches to evaluation: Phraseology and evaluative language*. Routledge, 2011.

Ken Hyland. Disciplinary discourses: Writer stance in research articles. *Writing: Texts, processes and practices*, 99121, 1999.

Ken Hyland. Authority and invisibility: Authorial identity in academic writing. *Journal of pragmatics*, 34(8):1091–1112, 2002.

Ken Hyland. Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, 7(2):173–192, 2005.

Timothy Jay and Kristin Janschewitz. The pragmatics of swearing. *Journal of Politeness Research. Language, Behaviour, Culture*, 4(2):267–288, 2008.

Elise Kärkkäinen. *Epistemic stance in English conversation: A description of its interactional functions, with a focus on I think*. John Benjamins Publishing, 2003.

Elise Kärkkäinen. Stance taking in conversation: From subjectivity to intersubjectivity. *Text & Talk-An Interdisciplinary Journal of Language, Discourse Communication Studies*, 26(6):699–731, 2006.

Elise Karkkainen. The role of i guess in conversational stancetaking. *Stancetaking in discourse, ed. Robert Englebretson*, pages 183–219, 2007.

Elise Kärkkäinen. Position and scope of epistemic phrases in planned and unplanned american english. In *New approaches to hedging*, pages 203–236. Brill, 2010.

Elise Kärkkäinen. I thought it was very interesting. conversational formats for taking a stance. *Journal of Pragmatics*, 44(15):2194–2210, 2012.

Tiina Keisanen. Stancetaking as an interactional activity: Challenging the prior speaker. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, pages 253–81, 2007.

Scott F Kiesling. Stance in context: Affect, alignment and investment in the analysis of stancetaking. In *iMean conference*, volume 15, 2011.

Paul Kockelman. Stance and subjectivity. *Journal of Linguistic Anthropology*, 14(2): 127–150, 2004.

Helga Koczogh. Scrutinizing the concept of (verbal) disagreement. *Argumentum*, 9: 211–222, 2013.

Irene Koshik. A conversation analytic study of yes/no questions which convey reversed polarity assertions. *Journal of Pragmatics*, 34(12):1851–1877, 2002.

Helga Kotthoff. Disagreement and concession in disputes: On the context sensitivity of preference structures. *LANGUAGE IN SOCIETY-LONDON-*, 22:193–193, 1993.

Karolina Krawczak. Epistemic stance predicates in english a quantitative corpus-driven study of subjectivity. *Subjectivity and Epistemicity Corpus, discourse, and literary approaches to stance*, page 303, 2014.

Sofia Lampropoulou and Greg Myers. Stance-taking in interviews from the qualidata archive. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, volume 14, 2012.

Jeanette Landgrebe. 'I think - you know'. Two epistemic stance markers and their significance in an innovation process. *Språk och interaktion 3*, page 107, 2012.

Andreas Langlotz and Miriam A Locher. Ways of communicating emotional stance in online disagreements. *Journal of Pragmatics*, 44(12):1591–1606, 2012.

Geoffrey N Leech. *Explorations in semantics and pragmatics*. John Benjamins Publishing, 1980.

Gene H Lerner and Celia Kitzinger. Extraction and aggregation in the repair of individual and collective self-reference. *Discourse Studies*, 9(4):526–557, 2007a.

Gene H Lerner and Celia Kitzinger. Introduction: person-reference in conversation analytic research. *Discourse Studies*, 9(4):427–432, 2007b.

Stephen C Levinson. Interrogative intimations: On a possible social economics of interrogatives. In J. P. de Ruiter, editor, *Questions: Formal, functional and interactional perspectives*, pages 11–32. Cambridge University Press, 2012.

Per Linell and Ivana Marková. Acts in discourse: From monological speech acts to dialogical inter-acts. *Journal for the Theory of Social Behaviour*, 1993.

Kris Liu and Jean E Fox Tree. Hedges enhance memory but inhibit retelling. *Psychonomic bulletin & review*, 19(5):892–898, 2012.

J. Lyons. *Semantics*. Semantics. Cambridge University Press, 1977.

J. Martin and P. White. *The Language of Evaluation: Appraisal in English*. Palgrave Macmillan UK, 2007.

James R Martin and Peter R White. *The language of evaluation*. Palgrave Macmillan, 2003.

Amita Misra and Marilyn A. Walker. Topic independent identification of agreement and disagreement in social media dialogue. *Evolution*, 460(460):920, 2013.

Jonathan Edward Mitchell. *The formal semantics of point of view*. PhD thesis, 1986. URL http://scholarworks.umass.edu/dissertations/AAI8701201/.

Peter Muntigl and William Turnbull. Conversational structure and facework in arguing. *Journal of Pragmatics*, 29(3):225–256, 1998.

Greg Myers. Stance-taking and public discussion in blogs. *Critical discourse studies*, 7(4):263–275, 2010.

Greg Myers and Sofia Lampropoulou. Impersonal you and stance-taking in social research interviews. *Journal of Pragmatics*, 44(10):1206–1218, 2012.

Sun-Young Oh. Overt reference to speaker and recipient in korean. *Discourse Studies*, 9(4):462–492, 2007.

Frank Robert Palmer. *Mood and modality*. Cambridge University Press, 2001.

Jung-ran Park. Linguistic politeness and face-work in computer-mediated communication, part 1: A theoretical framework. *Journal of the American Society for Information Science and Technology*, 59(13):2051–2059, 2008.

Nele Põldvere, Matteo Fuoli, and Carita Paradis. A study of dialogic expansion and contraction in spoken discourse using corpus and experimental techniques. *Corpora*, 2016.

Anita Pomerantz. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. Heritage Atkinson, J.M., editor, *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, Cambridge, 1984a.

Anita Pomerantz. Giving a source or basis: The practice in conversation of telling 'how i know'. *Journal of pragmatics*, 8(5):607–625, 1984b.

K. Precht. *Patterns of Stance in English*. Northern Arizona University, 2000.

Kristen Precht. Great versus lovely: Stance differences in american and british english. *Language and Computers*, 46(1):133–151, 2003a.

Kristen Precht. Stance moods in spoken english: Evidentiality and affect in british and american conversation. *Text & Talk-An Interdisciplinary Journal of Language, Discourse Communication Studies*, 23(2):239–258, 2003b.

Ken Prepin, Magalie Ochs, and Catherine Pelachaud. Beyond backchannels: co-construction of dyadic stancce by reciprocal reinforcement of smiles between virtual agents. In *International Conference CogSci (Annual Conference of the Cognitive Science Society)*, 2013.

Ellen F Prince, Joel Frader, Charles Bosk, et al. On hedging in physician-physician discourse. *Linguistics and the Professions*, 8:83–97, 1982.

Matthew Purver. SCoRE: A tool for searching the BNC. Technical Report TR-01-07, Department of Computer Science, King's College London, October 2001. URL http://www.dcs.kcl.ac.uk/technical-reports/papers/tr01-07.ps.gz.

Matthew Purver, Jonathan Ginzburg, and Patrick Healey. On the means for clarification in dialogue. In *Current and new directions in discourse and dialogue*, pages 235–255. Springer, 2003.

Randolph Quirk and David Crystal. *A comprehensive grammar of the English language*, volume 6. Cambridge Univ Press, 1985.

Victoria L Rubin. *Identifying certainty in texts*. PhD thesis, Syracuse University, 2006.

Harvey Sacks. On the preferences for agreement and contiguity in sequences in conversation. *Talk and social organization*, 54:69, 1987.

Harvey Sacks and Gail Jefferson. *Lectures on Conversation: Volumes I & II*. Number v. 1-2 in Lectures on conversation / Harvey Sacks. Ed. by Gail Jefferson. Blackwell, 1995.

Uli Sauerland and Penka Stateva. Scalar vs. epistemic vagueness: Evidence from approximators. In *Semantics and Linguistic Theory*, volume 17, pages 228–245, 2007.

Emanuel A Schegloff and Gene H Lerner. Beginning to respond: Well-prefaced responses to wh-questions. *Research on language and social interaction*, 42(2): 91–115, 2009.

Joanne Scheibman. *Point of view and grammar: Structural patterns of subjectivity in American English conversation*, volume 11. John Benjamins Publishing, 2002.

D. Schiffrin. *Discourse Markers*. Studies in Interactional Sociolinguistics. Cambridge University Press, 1988.

Deborah Schiffrin. Jewish argument as sociability. *Language in society*, 13(03): 311–335, 1984.

Suzanne Scott. Linguistic feature variation within disagreements: An empirical investigation. *Text & Talk-An Interdisciplinary Journal of Language, Discourse Communication Studies*, 22(2):301–328, 2002.

Jack Sidnell. "who knows best?": Evidentiality and epistemic asymmetry in conversation. *Pragmatics and Society*, 3(2):294–320, 2012.

Maria Sifianou. Disagreements, face and politeness. *Journal of Pragmatics*, 44(12): 1554–1564, 2012.

Maarit Siromaa. Resonance in conversational second stories: a dialogic resource for stance taking. 2012.

Sara W Smith. And people just you know like 'wow'discourse markers as negotiating strategies. *Discourse markers: Descriptions and theory*, 57:171, 1998.

Swapna Somasundaran and Janyce Wiebe. Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1-Volume 1*, pages 226–234. Association for Computational Linguistics, 2009.

Swapna Somasundaran and Janyce Wiebe. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124. Association for Computational Linguistics, 2010.

Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. Joint models of disagreement and stance in online debate. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, 2015.

Maria Stubbe and Janet Holmes. You know, eh and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of new zealand english. *Language & Communication*, 15(1):63–88, 1995.

Deborah Tannen. The pragmatics of cross-cultural communication. *Applied Linguistics*, 5(3):189–195, 1984.

Jean E Fox Tree and Josef C Schrock. Basic meanings of you know and i mean. *Journal of Pragmatics*, 34(6):727–747, 2002.

Arie Verhagen. *Constructions of intersubjectivity: Discourse, syntax, and cognition*. Oxford University Press on Demand, 2005.

M. Vihla. *Medical Writing: Modality in Focus*. Language and Computers: Studies in Practical Linguistics. Rodopi, 1999.

Diane Vincent and Laurent Perrin. On the narrative vs non-narrative functions of reported speech: A socio-pragmatic study. *Journal of Sociolinguistics*, 3(3): 291–313, 1999.

Marilyn A. Walker, Pranav Anand, Rob Abbott, Jean E. Fox Tree, Craig Martell, and Joseph King. That is your evidence?: Classifying stance in online political debate. *Decision Support Systems*, 53(4):719–729, 2012a.

Marilyn A Walker, Pranav Anand, Robert Abbott, and Ricky Grant. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 592–596. Association for Computational Linguistics, 2012b.

Marilyn A Walker, Jean E Fox Tree, Pranav Anand, Rob Abbott, and Joseph King. A corpus for research on deliberation and debate. In *LREC*, pages 812–817, 2012c.

Joseph B Walther. Interpersonal effects in computer-mediated interaction a relational perspective. *Communication research*, 19(1):52–90, 1992.

Joseph B Walther and Kyle P D'Addario. The impacts of emoticons on message interpretation in computer-mediated communication. *Social science computer review*, 19(3):324–347, 2001.

Shana Walton and Alexandra Jaffe. "stuff white people like": Stance, class, race, and internet commentary. *Digital discourse: Language in the new media*, pages 199–219, 2011.

Richard J Watts. *Politeness*. Cambridge University Press, 2003.

P White. The language of attitude, arguability and interpersonal positioning. *The Appraisal Website: Homepage. Disponível em: http://www. grammatics. com/appraisal/. Acesso em*, 14(06), 2004.

Peter Robert Rupert White. Appraisal-the language of evaluation and stance. 2002.

Peter RR White. Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text & Talk-An Interdisciplinary Journal of Language, Discourse Communication Studies*, 23(2):259–284, 2003.

Anna Wierzbicka. *English: Meaning and culture*. Oxford University Press, 2006.

Thomas Willett. A cross-linguistic survey of the grammaticization of evidentiality. *Studies in language*, 12(1):51–97, 1988.

# Appendix A

# Appendix 1

## A.1 Experiment Materials

# PARTICIPANT INSTRUCTION SHEET : BALLOON TASK

Please choose a nickname and then read ALL of this text before starting.

The task is to collaborate with your partner to resolve a dilemma. To do this you will be using:

(1) A character-by-character chat tool (the experimenter will explain how this works)

You use this window to communicate.

**The situation**

Three people are in a hot air balloon. The balloon is losing height and about to crash into the mountains. Having thrown everything imaginable out of the balloon, including food, sandbags and parachutes, their only hope is for one of them to jump to their certain death to give the balloon the extra height to clear the mountains and save the other two. But who is it to be?

The three people are:

**Dr. Nick Rivers** – a cancer research scientist who believes he is on the brink of discovering a cure for most common types of cancer. He is a good friend of Tom and Susie Derkins.

**Mrs. Susie Derkins** – a primary school teacher. She is over the moon because she is 7 months pregnant with her second child.

**Mr. Tom Derkins** – the balloon pilot. He is the husband of Susie, who he loves very much. He is also the only one with any balloon flying experience.

**Your Task**

You must discuss the pros and cons of keeping each of the three people in the balloon with your partner, and come to an agreement about who should be thrown out…

# Balloon Task Questionnaire

Please complete this form after you have finished the Balloon Task.

Title of Study:
**Investigation into the effects of a network-based chat tool on human dialogue & problem-solving**

Name used in chat tool ……

*Who did you agree should be thrown off the balloon? ……*

*Do you think this was the correct decision? ……yes/no*

*If no, who did you think it should be?……*

*How easy to come to an agreement did you find it? (please circle)*

| *Very easy* | | | *Neither easy nor difficult* | | | *Very difficult* |
|---|---|---|---|---|---|---|
| *1* | *2* | *3* | *4* | *5* | *6* | *7* |

*How easy did you think the discussion was to follow? (please circle)*

| *Very easy* | | | *Neither easy nor difficult* | | | *Very difficult* |
|---|---|---|---|---|---|---|
| *1* | *2* | *3* | *4* | *5* | *6* | *7* |

*Compared to a face-to-face conversation, how smoothly did you feel the conversation went? (please circle)*

| *Much more smoothly* | | | *The same* | | | *Much less smoothly* |
|---|---|---|---|---|---|---|
| *1* | *2* | *3* | *4* | *5* | *6* | *7* |

*Please enter any other comments you may have about the study below.*

……

……

……

……

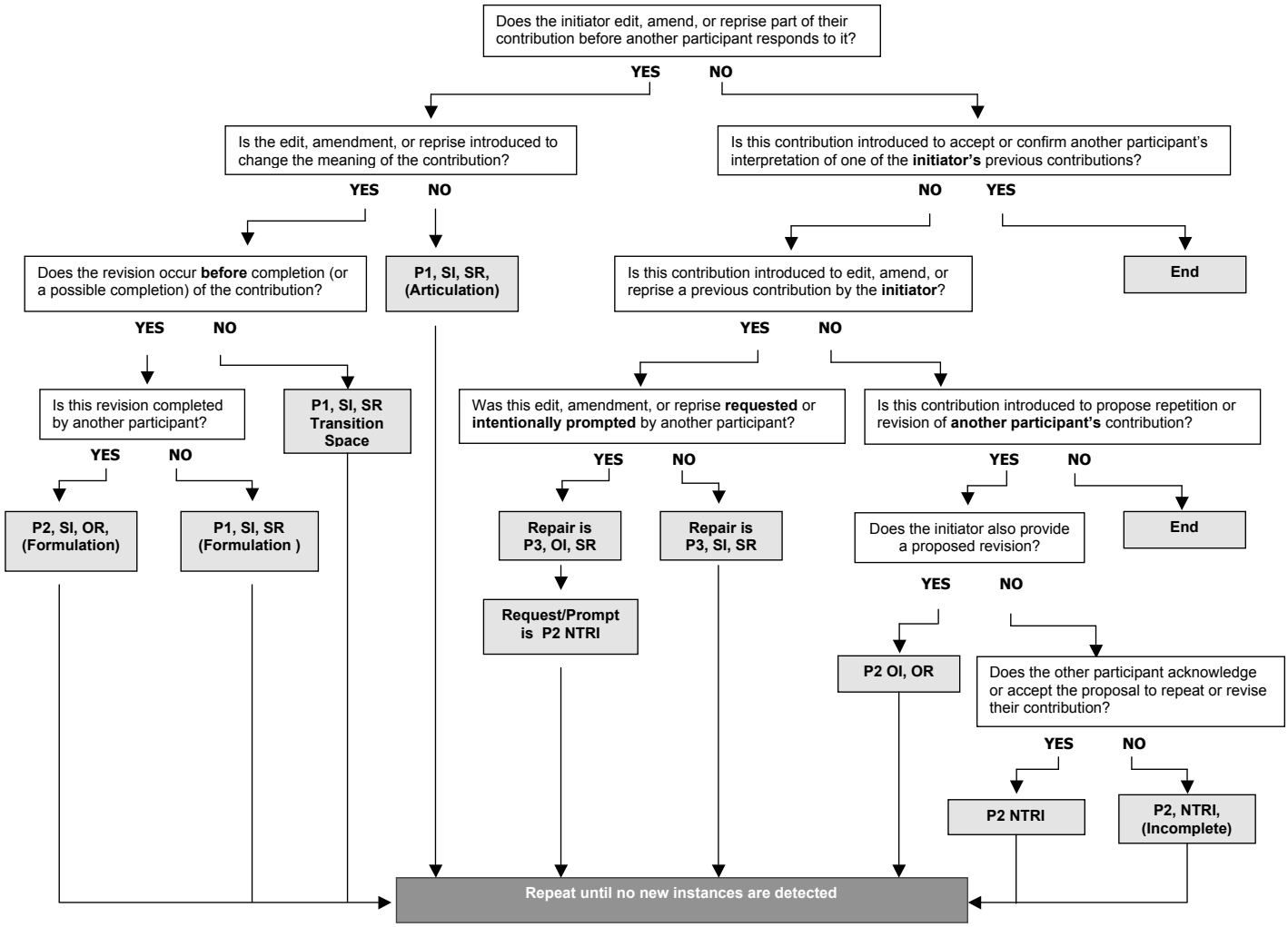
……

……

……

……

***Thank you for taking part in this study!***


Does the initiator edit, amend, or reprise part of their contribution before another participant responds to it?
YES
NO
Is the edit, amendment, or reprise introduced to change the meaning of the contribution?
Is this contribution introduced to accept or confirm another participant's interpretation of one of the initiator's previous contributions?
YES
NO
NO
YES
Does the revision occur before completion (or a possible completion) of the contribution?
P1, SI, SR, (Articulation)
Is this contribution introduced to edit, amend, or reprise a previous contribution by the initiator?
End
YES
NO
YES
NO
Is this revision completed by another participant?
P1, SI, SR Transition Space
Was this edit, amendment, or reprise requested or intentionally prompted by another participant?
Is this contribution introduced to propose repetition or revision of another participant's contribution?
YES
NO
YES
NO
YES
NO
P2, SI, OR, (Formulation)
P1, SI, SR (Formulation )
Repair is P3, OI, SR
Repair is P3, SI, SR
Does the initiator also provide a proposed revision?
End
Request/Prompt is P2 NTRI
YES
NO
P2 OI, OR
Does the other participant acknowledge or accept the proposal to repeat or revise their contribution?
YES
NO
P2 NTRI
P2, NTRI, (Incomplete)
Repeat until no new instances are detected

## A.2 Hedging and Boosting uses of 'I think' in the control conditions of experiment 1

| Turn | Label |
|---|---|
| I think because there's an element of risk with whether Nick will actually end up coming up with a cure for cancer | emphasis |
| I think Tom should definitely stay in the balloon | emphasis |
| I think Nick should definitely be the one to go | emphasis |
| i think we should stop now | emphasis |
| I think the doctor would be willing to sacrifice himself, as he's a good friend of theirs | emphasis |
| i think overall | unknown |
| I think the doctor would jump to his certain death | emphasis |
| i think its going to be pretty difficult to try and steer the balloon | emphasis |
| i think we have a couple mins left | hedge |
| i think we have made the right decision | hedge |
| i think | unknown |
| i think when they land the best 2 people for survival skills will be the 2 men | emphasis |
| i think the dr should jump | emphasis |
| i think its just us | emphasis/ambiguous* |
| I think thats actually gonna be the best idea | emphasis |
| i think nick | emphasis |
| i think so | hedge |
| i think what we said first | emphasis |
| i think tom | emphasis |
| i think making [h]er stay alive will be worse punishemnt | emphasis |
| i think dr nick would be good to keep because cancer is a huge problem in the world and no1 has found a cure for it yet and it would be good if he found it | emphasis |
| i think we did this in school | hedge |
| i think that too | emphasis |
| i think i'd go insane if i had to make that decision | emphasis |
| I think Tom wouldn't live with himself if Susie went. He loves her very much. | emphasis |

Table A.1 Coding of pragmatic effect of turn-initial 'I think' as hedge, emphasis or other

*Sequentially this is issued in response to the other participant asking if they think the researchers are reading what they are typing, and it read more as a dismissal of this; however, it is not possible to confirm this.

| Turn | Label |
|---|---|
| but *i think* we should think about the most valuable lives, not the quantity | emphasis |
| he might be but *i think* its unlikely that he is exaggerating | emphasis |
| whenever *i think* that nick should go, *i think* "Are susie and tom really that important?" | other |
| im still leaning towards nick staying o but *i think* we should finalise the decision | emphasis |
| as a dad *i think* he would want to see his wife and children live | emphasis |
| and even if Tom explained the basics to them, *I think* he'd be in a pretty panicky mood | emphasis |
| however in this otherwise justified decision *I think* thats alright | emphasis |
| yh basically *i think* the first person that needs to stay on the hot air balloon is mr tom | emphasis |
| I'm gonna fill in the sheet now, *I think* our time is up | hedge |
| the way *I think* of it is logistical who is going to be helped | other |
| okay *i think* thats the decision we would stick with | emphasis |
| ok honestly *i think* tom | emphasis |
| however *i think* if the cancer research scientist knew how to pilot the balloon, him and susie should stay on | emphasis |
| so *i think* nick shud go | emphasis |
| and *i think* the others could do that | hedge |
| but then again *i think* they'd prefer him then either one of them | emphasis |
| tbh *i think* no one should jump. they took the hot air balloon ride together | emphasis |
| but *i think* we should confirm | emphasis |
| No - pretty good *I think*. I saw a balloon crash land and everyone was ok, bit shaken. | hedge |

Table A.2 Coding of pragmatic effect of 'I think' (non-turn-initial) as hedge, emphasis or other

## A.3 Intercoder Reliability for Experiment 1

A second annotator was provided with the sub-sample of data and a set of written instructions. After reading the instructions and discussing for 10-15 minutes, I sat with the annotator and we talked through an example transcript to ensure that they felt comfortable with the task at hand and understood the instructions.

### A.3.1 Instructions for second coder

Hello and thank you for agreeing to annotate these three dialogue transcripts.

We are trying to capture the decision making process that the participants go through during the dialogue.

Therefore, your task is to label the current 'stance position' of each dialogue participant at the end of each turn (or row) in the dialogue transcript.

There is a column which corresponds to each participant's stance position, labelled with their name. Inside each column you can label the participant's stated position on whether the corresponding character should be killed (K), saved (S) or if they are undecided (u) about their fate.

If they strongly state that the wish to have a character die then the others are marked as saved. If they show some doubt then the stance position annotation changes to undecided. The stance position stays the same and carries over to the next row unless the text contradicts the stance position.

If more than one individual is up being considered for sacrifice then they would both be marked with a U for undecided. If they know that they want to save one person but are not agreed on who to kill then you would have S for saved in that character's column and U in the other two characters' columns.

Presume that each participant starts as undecided about the fate of the character, unless their first contribution clarifies otherwise. Continue through the spreadsheet updating each row based on what you have read in the text entry for that row.

Stance positions to chose from:

- Save Susie but undecided on who should die
- Save Nick but undecided on who should die
- Save Tom but but undecided on who should die
- Kill Tom
- Kill Nick
- Kill Susie

- Undecided

In the column marked *Stance position shift*, insert a '1' if on that line one of the participants' 'stance positions' has been updated, and '0' if it has not changed.

### A.3.2 Example of annotation

| ParticipantID | Sender | Text | Stance state | | |
|---|---|---|---|---|---|
| | | | Tom | Nick | Susie |
| 2 | P1 | I think we should kill off tom | k | s | s |
| 1 | P2 | yeah | k | s | s |
| 1 | P2 | poss | u | u | u |
| 1 | P2 | But Nick and Susie can't fly it | s | u | u |
| 1 | P2 | so they'll end up dying anyway | s | u | u |
| 2 | P1 | Yeah but Nick is a doctor so he can work out how to fly a balloon surely? | k | s | s |
| 1 | P2 | no ah, yes. | u | s | u |
| 1 | P2 | Definitely not Nick to die though | u | s | u |
| 2 | P1 | Yeah needs of the many outway the needs of the few | k | s | s |
| 1 | P2 | he'll save many in the future, if he's not all chat | u | s | u |
| 2 | P1 | looooooool | k | s | s |
| 1 | P2 | I agree | s | s | k |
| 1 | P2 | I disagree also, Susie.. | s | s | k |
| 2 | P1 | no susie has a little waine on the way so we need to keep her alive | k | s | s |
| 1 | P2 | Yes, d-turbo | s | s | k |
| 1 | P2 | although, some would say that the child is not alive yet and therefore wouldn't be taking a life | s | s | k |

Fig. A.1 Example annotation

Figure A.1 shows an excerpt from the example annotated transcript provided to the annotator.

## A.4 Intercoder Reliability for Experiment 2

A second annotator was provided with the sub-sample of data and a set of written instructions. After reading the instructions and discussing for 10-15 minutes, I sat with the annotator and we talked through an example transcript to ensure that they felt comfortable with the task at hand and understood the instructions.

### A.4.1 Instructions for second coder

In the column marked *Stance position* enter an annotation for each row based on the corresponding message data. Based only on what is written in the text, for each

line of dialogue select the appropriate annotation which conveys who the particiapnt has stated should be killed. This need not necessarily convince you that they have truly changed their opinion on the matter as this becomes complicated to decipher, but judged on the message that they have typed, if it implies an alternative person is being selected (even if only temporarily) to be sacrificed, note this down as your annotation. The annotation should only record who the participant states should be killed, and if this is not yet clear, or they have shown that they are undecided again then you mark their stance status as undecided.

Stance positions to chose from:

- Kill Tom
- Kill Nick
- Kill Susie
- Undecided

In the column marked *Stance position shift*, insert a '1' if on that line one of the participants' 'stance positions' has been updated, and '0' if it has not changed.

### A.4.2 Example of annotation

| Name | text | P1 | P2 |
|---|---|---|---|
| P2 | but it will just miss the mountains | uuu | uuu |
| P1 | yeah but when do you have alone mountain? | uuu | uuu |
| P2 | who knows | uuu | uuu |
| P1 | lol | uuu | uuu |
| P2 | ok | uuu | uuu |
| P2 | so whta are your first thoughts? | uuu | uuu |
| P2 | I personally say chuck doctor out | uuu | Nick |
| P2 | haha | uuu | Nick |
| P1 | so we need to decide lol | uuu | Nick |
| P1 | I would say the woman | Susie | Nick |
| P1 | 2 lifes for generations | Susie | Nick |
| P2 | but she has a tiny baby in her belly!!!!!!!! | Susie | Nick |
| P1 | worlds full | Susie | Nick |
| P2 | hahahaha omh | Susie | Nick |
| P2 | *omg | Susie | Nick |
| P2 | no you can't | Susie | Nick |
| P2 | her husband would have to basically kill his wife and his future baby | Susie | Nick |
| P1 | also the doctor says he 'believes' he is on the brink | Susie | Nick |
| P1 | doesnt mean he is doing well tho | Nick | Nick |
| P2 | exactly | Nick | Nick |
| P2 | so get rid of him | Nick | Nick |
| P2 | haha | Nick | Nick |
| P1 | but realistically if you consider the value should the man/husband/dad jump and provide a quick lesson to the doctor on how to drive the balloon | Tom | Nick |
| P1 | as he is clearly smart | Tom | Nick |
| P2 | yeah | Tom | Nick |
| P2 | that's an option | Tom | Nick |

Fig. A.2 Example annotation

Figure A.2 shows an excerpt from the example annotated transcript provided to the annotator.

# Appendix B

# Appendix 2

## B.1 Ethical Considerations

A Research Ethics Questionnaire was submitted and approved for the studies contained within this thesis. The letter confirming this is included here, together with the Participant Instructions and Consent form that were presented to all participants who took part in the experiments.

**Queen Mary, University of London**
Room W117
Queen's Building
Mile End Road
London
E1 4NS

**Queen Mary Research Ethics Committee**
Hazel Covill
Research Ethics Committee Administrator
Tel: +44 (0) 20 7882 7915
Email: h.covill@qmul.ac.uk

c/o Professor Pat Healey
CS 410
Department of Computer Science
Queen Mary University of London
Mile End
London

29th September 2014

To Whom It May Concern:

**Re: QMREC1379a – Investigation into the Effects of a Chat Tool on Human Dialogue and Discussion Style.**

I can confirm that Ms Shauna Concannon has completed a Research Ethics Questionnaire with regard to the above research.

The result of which was the conclusion that her proposed work does not present any ethical concerns; is extremely low risk; and thus does not require the scrutiny of the full Research Ethics Committee.

Yours faithfully

Ms Hazel Covill
Research Ethics Committee Administrator

## 18. CONSENT FORM (guideline 18)

Please complete this form after you have read the Information Sheet and listened to an explanation about the research.

Title of Study:
**Investigation into the Effects of a Network-based Chat Tool on Human Dialogue and Problem-Solving**

Queen Mary Research Ethics Committee Ref: _____

- Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.
- If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

***Please cross out as necessary***

*Have you been asked to consent for yourself or on behalf of someone else?.....self/other*

*Have you read the Information Sheet (Appendix A)?.....yes/no*

*Have you had an opportunity to ask questions and discuss this study?.....yes/no*

*Who have you spoken to?.....*

*Do you understand that you are free to withdraw from the study at any time without having to give a reason for withdrawing and will receive payment in full for participating?.....yes/no*

*Do you understand that everything you type using the chat tool will be recorded (There is no audio recording).....yes/no*

*Do you understand that you are free to decide after the task is completed whether you are happy for the recorded conversation to be studied and stored?.....yes/no*

*Do you have some experience of using text-based chat tools such as MSN Messenger or ICQ?.....yes/no*

*Have you declared your involvement in other research studies currently under way or undertaken in the last 12 months?.....yes/no*

*Do you consent to the processing of your personal information for the purposes of this research study with the understanding that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998?.....yes/no*

*Do you agree to be contacted in the future by QMUL researchers who would like to invite you to participate in follow up studies to this project or in future studies of a similar nature?.....yes/no*

*Do you understand that at any point after the experiment you are entitled to a "debriefing" to discuss any concerns that might have arisen due to having taken part in this study.....yes/no*

**Participant's Statement:**

I _____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

**Investigator's Statement:**

I _____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer.

Signed: Date:

## 17. INFORMATION SHEET (guideline 17)

REC Protocol Number……

**YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET**

### Investigation into the Effects of a Network-based Chat Tool on Human Dialogue and Problem-Solving

We would like to invite you to participate in this original research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Your decision will not affect anything concerning your studies at Queen Mary. Before you decide whether you want to take part, it is important for you to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. If you do decide to take part, please let us know beforehand if you have been involved in any other study during the last year.

We are investigating the usability of a network-based chat program and maze game in order to study their effect on the way people interact with each other, and would like to ask you to participate in a study using these two programs. You will have two tasks. One will involve communicating via the chat program, and the other will be to solve a series of maze puzzles. In total, both tasks should take less than two hours. Afterwards we welcome any feedback about your experience using the programs.

The conversation will be recorded and kept for study. The records will be anonymous: when you begin the task you will be asked to choose a nickname for the chat, and your real name will not be stored in the records.

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason. In particular, after the tasks are complete you will be able to decide whether you are happy for the recorded conversation to be studied and stored.

In the event of you suffering any adverse effects as a consequence of your participation in this study, you will be compensated through Queen Mary University of London's 'No Fault Compensation Scheme'.